CESifo

AREA CONFERENCES 2020

Public Economics

Online, 8–9 December 2020

Welfare Effects of Pension Reforms

Andreas Haller

CESifo

Welfare Effects of Pension Reforms*

Andreas Haller
November 19, 2019

Click here for the most recent version

Abstract

In almost all developed countries, policy makers have implemented pension reforms by increasing statutory retirement ages, lowering pension levels and/or adjusting pension formulas to address demographic change. This paper provides a novel, unifying framework to evaluate the welfare effects of such pension reforms. I show that the welfare effects of any reform rest crucially on the "behavioral fiscal multiplier"–the total fiscal effect relative to the mechanical fiscal effect (the mechanical effect is the fiscal effect absent any behavioral responses). Behavioral fiscal multipliers can be readily estimated with reduced-form methods using data on contributions to and transfers from the entire welfare state system. To illustrate my framework, I exploit a series of pension reforms in Austria. I find that increasing the early retirement age has a behavioral multiplier of 1. This means that the total fiscal effect is purely mechanical and there is no fiscal effect from behavioral adjustments. In response to the reform workers spend more time in employment, which generates additional social security contributions. However, individuals also spend more time in unemployment, which generates additional expenditures on unemployment insurance benefits. These two effects cancel, leading to a net-zero fiscal effect of behavioral responses. This finding implies that increasing the Austrian early retirement age is not welfare-enhancing–unless one thinks that $1 in the hands of an early retiree has a lower social value than $1 in public funds. By contrast, reducing pension levels generates a multiplier of 1.5. This policy induces some workers to stay longer in employment without triggering substitution to other welfare benefits. As a result, reducing pension levels is welfare improving, provided that taking $1 away from a retiree is associated with a social loss smaller than $1.5. In a standard calibration of the model, the social loss is smaller than $1.5 for reasonable values of risk aversion suggesting that reducing pension levels was welfare-improving. My framework can also rank the welfare effects of the two reforms. Based on my estimates, a social planner with preferences for redistribution clearly favors reducing pension levels over increasing the early retirement age.

*Address: Andreas Haller, University of Zurich, CH-8000 Zürich, Switzerland; email *andreas.haller@uzh.ch*.
I am grateful to my advisors Josef Zweimüller and Florian Scheuer for their guidance and support. I thank Doug Bernheim, Pietro Biroli, Mark Duggan, Liran Einav, Giacomin Favre, Sandro Favre, Lucija Muehlenbachs, Petra Persson, Hannes Schwandt, Isaac Sorkin, Stefan Staubli, Danny Yagan and seminar participants at the University of Zurich, Stanford, Calgary, Max Planck and EALE for helpful comments.

1 Introduction

Due to dramatically aging populations throughout the world, pension reforms are on the political agenda everywhere. The basic idea of pension reforms is to incentivize workers to delay their retirement, thus increasing the ratio of workers to pensioners and easing the demographic burden on pay-as-you-go social security systems.[1] Many countries have implemented pension reforms in the recent past. For instance, the U.S. has started to gradually increase the full retirement age from 65 to 67 and almost all European countries have implemented increases in the statutory retirement ages. Many countries have increased work incentives by reducing pension levels and/or by increasing the actuarial fairness of pensions. The hope is that these measures induce workers to postpone their retirement. While many reforms have already been implemented, future reforms are inevitable to avoid the financial collapse of pay-as-you-go pension systems. The key question is: How should we reform pension systems?

To answer this question we need to understand the welfare effects of pension reforms. While there exists a rich literature that evaluates the labor supply and fiscal effects of pension reforms (e.g. Mastrobuoni 2009; Behaghel and Blau 2012; Staubli and Zweimüller 2013; Manoli and Weber 2016a; Cribb et al. 2016), the link between the empirical estimates from the program evaluation literature to welfare consequences is missing so far. This paper aims to provide this link. I show that the welfare effect of a pension reform crucially depends on the "behavioral fiscal multiplier" and that behavioral fiscal multipliers can be credibly estimated with reduced-form methods. The behavioral multiplier of a reform is defined as the total fiscal effect divided by the mechanical fiscal effect. The mechanical fiscal effect measures the hypothetical effect of a reform on fiscal revenue if individual behavior did not respond to the reform. The mechanical fiscal effect can be directly calculated holding pre-reform behavior fixed. The total fiscal effect of a pension reform can be estimated with program evaluation methods. The existing literature estimates the total fiscal effects of pension reforms but usually does not identify the mechanical fiscal effect. However, this extra step of relating the total effect to the mechanical effect (i.e. the behavioral multiplier) is crucial for welfare evaluation.

Why is the behavioral multiplier key for welfare evaluation? One can think of a pension reform as taking away one dollar from a retiree and transferring it to public funds. Different pension reforms simply target different groups of retirees. For example, an increase in the early retirement

[1] In pay-as-you-go pension systems, working-age individuals pay for the pensions of the currently old generation.

age takes away one dollar from early retirees, a cut in pension levels takes away one dollar from all retirees. A reform then mechanically increases public funds by one dollar. Additionally, a reform induces behavioral responses. Individuals adjust to the new rules, e.g. by delaying retirement and working longer, which further increases public funds. In total, taking away one dollar from a retiree increases public funds by the behavioral multiplier. However, taking away one dollar from a retiree comes at a social cost. This social cost depends on the relative social valuation of one dollar in the hands of a retiree versus one dollar in public funds. I refer to this cost as "the social value of the dollar." If one dollar is socially more valuable in the hands of a retiree, the social value of the dollar is above one. If one dollar is socially more valuable in public funds, the social value of the dollar is below one. The social value of the dollar, therefore, measures the social cost of transferring one dollar from retirees to public funds in dollar terms. More specifically, the social value of the dollar depends on three things. First, it depends on how valuable this dollar is to the affected retirees (measured by their marginal utility of consumption). Second, it depends on the welfare weights the social planner attaches to the group of affected retirees. Third, it depends on the social value of public funds, i.e. the social value of whatever else the planner would use this dollar for, or put differently, the social cost of raising public funds.[2]

In summary, taking away one dollar from retirees (a pension reform) has a social benefit of increasing public funds by the behavioral multiplier, but comes at the cost of the social value of the dollar. Hence, to evaluate the welfare effect of a reform, we need to compare the behavioral multiplier to the social value of the dollar (social benefits vs. social costs). If the behavioral multiplier is larger than the social value of the dollar, the reform is welfare improving (social benefits exceed social costs) and vice versa. The social value of the dollar does not directly relate to moments in the data.[3] After all, the social value of the dollar depends on the value that society as whole puts on a marginal increase of public funds relative to a decrease in pension generosity. This is a judgment call but under reasonable assumptions on the efficiency of the tax system and social preferences the social value of the dollar is larger than one, i.e. one dollar in the hands of a retiree has at least a social value of one dollar. In contrast, the behavioral multiplier can be readily estimated from data on fiscal costs and revenues that are generated by individuals' behavioral responses to the reform. The behavioral multiplier is essential, because it is

[2]Importantly, economic theory tells us that behavioral adjustments to small reforms do not have direct welfare effects (by the envelope theorem) and hence do not enter in the social value of the dollar.

[3]I am not aware of other papers that have related the social value of a dollar to empirical estimates in the retirement context. In the unemployment insurance literature, there are exciting new approaches on this frontier (e.g. Hendren (2017) and Landais and Spinnewijn (2019)). These approaches might also be applied in the retirement context but require data on consumption and saving decisions of individuals.

the benchmark against which to judge the social value of the dollar. Under certain circumstances, knowing only the behavioral multiplier (but not the social value of the dollar) suffices to assess whether a pension reform was welfare improving or not. In my empirical analysis, the size of the estimated multipliers (and the characteristics of the affected workers) imply that only mild assumptions on social preferences are needed to make clear-cut statements about the welfare effects of these reforms.

In the empirical analysis of the paper, I estimate the behavioral multiplier of increasing the early retirement age and the behavioral multiplier of reducing pension levels by exploiting a series of pension reforms in Austria. Austria provides an ideal set up for three main reasons. First, seven pension reforms changed the early retirement age, pension levels and actuarial fairness of the pension formula at least once. Austrian policy makers phased in most policy changes yielding vast quasi-experimental variation. Second, the Austrian pension formula is very similar to other developed countries' pay-as-you-go formulas and the Austrian pension reforms changed margins that relate to current policy discussions. Third, the Austrian Social Security Data (ASSD) is an ideal data source to estimate behavioral multipliers of pension reforms. The ASSD does not only include information on the complete labor market and earnings history of workers since 1972 but also the history of take-up of welfare state programs (such as unemployment insurance, disability insurance and sickness benefits). As I will demonstrate below, changes in take-up of social insurance programs crucially affect the size of the behavioral multiplier.

To estimate the behavioral multiplier of increasing the early retirement age (ERA), I exploit two pension reforms in 2000 and 2003 that increased the ERA in steps from 55 to 60 for women and from 60 to 65 for men. The increase in the ERA is phased in by quarters of birth. With a difference-in-difference strategy, I find large and positive fiscal revenue effects of increasing the ERA. However, this fiscal revenue effect is purely mechanical, implying that the behavioral multiplier of this reform is one and in some cases as low as 0.9. This is surprising. Since individuals can no longer retire early, we expect them to work longer and pay more taxes, leading to additional fiscal revenue and hence a large behavioral multiplier. While there is a positive fiscal effect through additional payroll tax revenue, there are also additional expenditures from individuals substituting to unemployment benefits. It turns out that these two effects cancel out each other. The net fiscal effect from behavioral adjustments is therefore zero, leading to a behavioral multiplier of one. Interestingly, this finding implies that increasing the Austrian early retirement age is not welfare-enhancing – unless one thinks that $1 in the hands of an early retiree has a lower social value than \$1 in public funds. As I show in the theoretical part of the paper, a natural lower bound for the social value of the dollar is one. Hence, the ERA reform with a behavioral multiplier of one is unlikely to be welfare-enhancing.

To estimate the behavioral multiplier of reducing pension levels, I exploit a 1988 reform using a regression discontinuity design. The reform changed the pension formula by date of birth in a particular way that made pensions by about 1.25% less generous on average.[4] For this reform, I find large behavioral multipliers of around 1.5. In response to the reform, I find that individuals delay their retirement and work longer. This increases fiscal revenue both by a reduction in spending on pensions and by additional pay-roll tax revenue. There is no substitution to other welfare benefits that would counteract this effect. In sum, this produces a behavioral multiplier of 1.5. Hence, reducing pension levels is welfare improving – provided that taking one dollar away from a retiree is associated with a social loss smaller than 1.5 dollars. In a standard parametrization of the model with CRRA utility I find that the social loss is smaller than 1.5 dollars for reasonable values of risk aversion. Hence, reducing pension levels is welfare-enhancing. The result that reducing pension levels generates behavioral multipliers substantially larger than 1 is surprising. A reduction in pension levels has a large mechanical effect, since individual's pensions are lower for the rest of their life. Moreover, a change in pension levels might be less salient and more complicated to understand than increasing the ERA. Therefore, one might not expect that individuals change their labor supply and retirement decision in response to this reform.

My framework also allows comparing the welfare effect of increasing the ERA to the welfare effect of reducing pension levels. Based on my estimates, a social planner with preferences for redistribution clearly favors reducing pension levels over increasing the ERA. Reducing pension levels has a substantially higher behavioral multiplier of 1.5 compared to the ERA reform with a behavioral multiplier of 1. Moreover, the pension level reform did not uniformly cut pension levels but disproportionately affected high income earners, while the ERA reform affected individuals across the income distribution. This implies a lower social value of the dollar for the pension level reform, since taking away one dollar from high income retirees is less costly than taking away one dollar from all retirees. Hence, reducing pension levels has a higher social benefit (behavioral multiplier) and comes at lower social costs (social value of the dollar) compared to increasing the ERA. Therefore, reducing pension levels is preferable to increasing the ERA in the Austrian

[4] An individual's pension is determined by her assessment basis multiplied with her pension coefficient. The assessment basis measures the average earnings over a specific period (assessment period) after applying a cap to earnings in each year. The 1988 pension reform increases the assessment period from the last 10 years to the last 11 years for men born after January 1st 1928 and for women born after January 1st 1933. Due to seniority wages this decreases pension by 1.25% on average.

context.

The more general message of my analysis is that spillovers to other welfare systems are of first-order importance and labor market opportunities of older workers are key in this respect. The behavioral multiplier summarizes these effects. If my framework is applied to reforms in other countries, the welfare implications can be very different depending on the labor market responses of affected workers.

Contribution to Literature. The logic that the behavioral multiplier is central to welfare evaluation is not new. This is a direct consequence of the envelope theorem and the foundation of the sufficient statistics literature. The previous literature on sufficient statistics has mainly focused on optimal unemployment insurance and taxation. There is a lot of recent and exciting work done in these areas (for an overview of this literature see Chetty and Finkelstein (2013); Kleven (2018) and for recent papers on unemployment insurance see Chetty (2006a); Shimer and Werning (2007); Chetty (2008); Schmieder et al. (2012); Kolsrud et al. (2018)). This is the first paper to take this idea to the retirement context and show that reduced-form estimates are informative for welfare effects of pension reforms. In the retirement context, with complicated dynamics and multiple generations, it is not obvious what the relevant reduced-form estimates are and whether we can estimate them. This paper demonstrates that the behavioral multiplier of a reform is central and that it can be credibly estimated with reduced-form methods. There are two recent working papers that closely connect to my paper. Lee et al. (2019) estimate the fiscal externality (behavioral multiplier minus 1) for two unemployment policies. Hendren and Sprung-Keyser (2019) estimate the marginal value of public funds (willingness to pay for a policy divided by the behavioral multiplier) for 133 historical policy changes in the United States, focusing on policies in social insurance, education and job training, taxes and cash transfers, and in-kind transfers. Both papers argue, as I do as well, that these measures a key for welfare analysis of policy changes.

The empirical part of this paper contributes to the large and growing reduced-form literature evaluating past pension reforms (Duggan et al., 2007; Mastrobuoni, 2009; Behaghel and Blau, 2012; Staubli and Zweimüller, 2013; Manoli and Weber, 2016a; Cribb et al., 2016; Seibold, 2019) by evaluating the labor supply and fiscal implications of increasing the ERA and reducing pension levels. Staubli and Zweimüller (2013) and Manoli and Weber (2016b) study the same ERA reform as this paper. My analysis uses a slightly different identification strategy and estimates the behavioral multiplier of the reform as a new outcome. The theoretical framework provides a new perspective on the welfare effects of this reform. The pension level reform has not been studied before.

There is also a large literature on structural retirement models (Gustman and Steinmeier, 1986; Stock and Wise, 1990; Berkovec and Stern, 1991; Rust and Phelan, 1997; French, 2005; van der Klaauw and Wolpin, 2008; Iskhakov, 2010; Laitner and Silverman, 2012; Imrohoroglu and Kitao, 2012; Gustman and Steinmeier, 2015). This literature performs counterfactual analysis of social security reforms in estimated models to understand the labor supply and welfare effects of pension reforms. I view this structural approach as complementary to my paper. For instance, in my empirical analysis I cannot observe the labor supply response to a reform over the entire life-cycle of an individual (especially at younger ages). In policy simulations, French (2005) finds very little life-cycle variation in hours worked between ages 30 and 55, suggesting that my empirical analysis might not be missing a substantial effect by ignoring outcomes at younger ages. In the other direction, careful reduced-form empirical analysis of pension reforms can be informative for structural model building.

Road-map. Section 2 develops the theoretical framework for welfare evaluation. Section 3 describes the institutional background in Austria, the seven pension reforms and the data. The empirical analysis in section 4 is divided in two parts. Section 4.1 estimates the behavioral multiplier of increasing the early retirement age. Section 4.2 estimates the behavioral multiplier of decreasing pension levels. Finally, Section 5 concludes. Various model extensions and robustness checks are relegated to the Appendix.

2 Model

This section develops the framework to evaluate welfare effects of pension reforms. The model is in the spirit of the sufficient-statistics literature. I show that in a large class of models the multiplier of a pension reform, which can be estimated with reduced-form methods, is key for welfare evaluation.[5] Pension reforms are challenging to evaluate because there are effects within and between generations. I start with a life-cycle model of only one generation to illustrate the key trade-offs. In Appendix A.4 I show that with multiple overlapping generations we want to estimate the exact same multiplier as with only one generation. The logic of the argument is that we want to design the optimal pension formula for each generation and hence need to know the multiplier of a reform within one generation.

[5] In what follows, I will use "multiplier" when I refer to the "behavioral multiplier."

Agents. There is a continuum of agents indexed by $i \in I$ facing a life-cycle of T periods. Agent's expected life-time utility, denoted by $U_i(C, \Pi, X)$, depends on their consumption C, other choices Π (such as labor supply), and the full state history X. In each period, an agent is in a specific state x_t with a state history $x^t = \{x_i\}_{i=0}^{t-1}$. A state is described by assets a_t and a vector of other states s_t, i.e. $x_t = \{a_t, s_t\}$. The other states s_t can include, but are not limited to, labor market status, past earnings, health, productivity and mortality. This formulation is flexible, allowing, for instance, for utility depending on labor markets status, health status, productivity or age. In particular, this can also capture differences in longevity. While I fix the maximum length of life to T periods, this is not restrictive. Differences in longevity can be modeled as a state variable "death" that sets the utility function to zero for all future periods and T can be arbitrarily large. Each period, an agent chooses consumption $c_t(x^t)$ and other choices $\pi_t(x^t)$, which includes labor income $y_t(x^t)$ and a vector $p_t(x^t)$ of other behaviors. $C \equiv \left(c_0(x^0), c_1(x^1), \ldots, c_{T-1}(x^{T-1})\right)$ and $\Pi \equiv \left(\pi_0(x^0), \pi_1(x^1), \ldots, \pi_{T-1}(x^{T-1})\right)$ denote the contingent life-cycle plan of an agent. The agent therefore solves the following optimization problem

$$V_i = \max_{C,\Pi} U_i(C, \Pi, X) \tag{1}$$

subject to the constraints

$$a_{t+1} = (1 + r_t)a_t + y_t(x^t) - \tau(x^t) + b(x^t) - c_t(x^t) - q(p_t(x^t)) \tag{2}$$

and

$$s_{t+1} = f\left[s^t, \pi_t(x^t), \varepsilon_t\right]. \tag{3}$$

The constraints impose minimal structure on the evolution of state variables. For assets, I impose the standard budget constraint (2) stating that for every state history x^t in each period t assets tomorrow a_{t+1} equal assets today a_t plus capital income $r_t a_t$, plus after tax income $y_t(x^t) - \tau(x^t)$ and benefit receipt $b(x^t)$, minus consumption $c_t(x^t)$ and expenditures of the other choices $q(p(x^t))$.[6] Constraint (3) characterizes the evolution of the other state variables such as labor market status or health, and is assumed to follow some process $f\left[s^t, \pi_t(x^t), \varepsilon_t\right]$. This process is allowed to depend on the full state history s^t, the random component ε_t and choices π_t.

[6] $q(\cdot)$ is mapping the other choices $p_t(x^t)$ into monetary expenditures.

Social Planner. The social planner can specify state-history-dependent transfers $b(x^t)$ and taxes $\tau(x^t)$. These functions capture all kinds of (un-)conditional transfers and taxes and therefore, in principle, describe the entire welfare state. When choosing the benefit and tax function, the social planner faces a revenue constraint $G(b,\tau) \geq \bar{G}$, where revenue is defined as the present value of taxes minus transfers. The planner's objective function sums over the indirect utilities V_i of all agents weighted by welfare weights ω_i. The planner therefore solves

$$\max_{b(\cdot),\tau(\cdot)} W(b,\tau) = \int \omega_i V_i di \tag{4}$$

subject to

$$G(b,\tau) = \sum_{t=0}^{T-1} [1+r_t]^{-t} E\left[\tau(x^t) - b(x^t)\right] \geq \bar{G}. \tag{5}$$

The expectation operator in (5) is defined over the distribution of state histories, which depends on the agent's choices C and Π.[7]

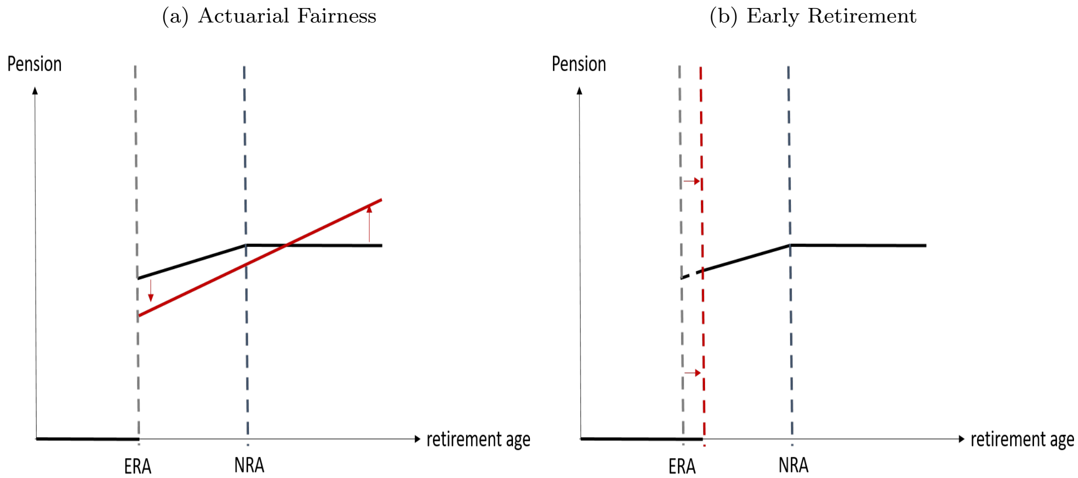
Fréchet Derivative. The social planner is optimizing over functions. To formalize this I use the Fréchet derivative, which is a generalization of the concept of directional derivatives to functions.[8] The Fréchet derivative $\delta W(b;h)$ measures the change in social welfare $W(b,\tau)$ if the benefit function $b(\cdot)$ is tilted in direction of the function $h(\cdot)$. Intuitively, one can think of changing the current benefit function $b(\cdot)$ in direction $h(\cdot)$ as a specific reform that changes the pension formula in particular way. Different directions $h(\cdot)$ can capture changes in pension generosity (level shift in benefit function), in actuarial fairness (slope of the benefit function) and early and normal retirement age (shift to the right in the age dimension). Figure 1 illustrates this idea graphically. In summary, the Fréchet derivative is a powerful tool to capture arbitrarily complicated (marginal) changes in the benefit function and therefore any small pension reform can be evaluated in my framework.

Optimal Benefit Function. Suppose the planner implements a particular pension reform, i.e. the benefit function is changed in the direction of a measurable function $h : X \to \mathbb{R}$. The welfare effect of this reform is given by

[7] Formally, $\sum_{t=0}^{T-1} [1+r]^{-t} E\left[\tau(x^t) - b(x^t)\right] = \int \sum_{t=0}^{T-1} [1+r]^{-t} \left[\tau(x^t) - b(x^t)\right] \mu(d(\varepsilon_0,\ldots,\varepsilon_{T-1})|x_0,C,\Pi)$ where $\mu(d(\varepsilon_0,\ldots,\varepsilon_{T-1})|x_0,C,\Pi)$ is the product measure on the cross product of Υ^T (the space of the error terms ε_t) and T copies of Υ, since $x_t = \{a_t, s_t\}$ can be expressed in terms of $(\varepsilon_0,\ldots,\varepsilon_{T-1})$ via (2) and (3). For notational ease, I assume that the distribution of state histories has a density function, denoted by $\mu(x^t;C,\Pi)$. This is not crucial for the argument but allows me to write the expectation operator as $\sum_{t=0}^{T-1} \int [1+r]^{-t} \left[\tau(x^t) - b(x^t)\right] \mu(x^t;C,\Pi)dx^t$.

[8] See Luenberger (1997) for a formal treatment.

Figure 1: Illustration Fréchet Derivative



Notes: This figure illustrates the idea of the Fréchet derivative. Suppose the black line is the pre-reform benefit function and the red line the post-reform benefit function. In panel (a) we change the slope of the benefit function. Individuals who retire early face a reduction in their benefits. Hence, the direction of change $h(\cdot)$ is negative for these individuals. Individuals who retire later face an increase in their benefits and their direction of change $h(\cdot)$ is positive. Panel (b) illustrates an increase of the early retirement age (ERA). This shifts the whole benefit function to the right.

$$\delta W(b;h) = \int \omega_i E_i \left[\sum_{t=0}^{T-1} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)} h(x^t) \right] di + \lambda * \delta G(b;h). \tag{6}$$

The first term measures the direct welfare effect of changing the benefit function. This direct welfare effect is given by the change of the benefit function $h(x^t)$ multiplied by the marginal utility of consumption $\frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)}$ of agent i in that state multiplied by her welfare weight ω_i. This direct effect only depends on the mechanical effect of the reform. The Envelope theorem ensures that behavioral adjustments do not have a first-order welfare effect. The second term, $\lambda * \delta G(b;h)$, captures the social value of the fiscal effect of the reform. $\delta G(b;h)$ is the total change in fiscal revenue and consists of two separate effects: The mechanical and behavioral fiscal effect. The mechanical fiscal effect, $M(h)$, is the hypothetical change in fiscal revenue if agents' behavior did not respond to the reform. Formally, this simply sums up all changes in the benefit function $h(\cdot)$ at the pre-reform behavior, i.e. at the pre-reform state-history distribution

$$M(h) = \int \sum_{t=0}^{T-1} [1+r_t]^{-t} E_i \left[-h(x^t) \right] di. \tag{7}$$

The behavioral effect $B(h)$ measures the hypothetical change in fiscal revenue of the reform if only agents' behavior adjusted and the benefit function was held constant. Formally, the behavioral effect is given by

$$B(h) = \int \sum_{t=0}^{T-1} \int [1+r]^{-t} \left[\tau(x^t) - b(x^t)\right] \delta\mu(x^t; h) dx^t di, \tag{8}$$

where $\delta\mu(x^t; h)$ measures the change in the state-history distribution due to changes in the agents' behavior. The sum of these two effects yields the total fiscal revenue effect $\delta G(b; h) = M(h) + B(h)$. Lastly, the multiplier on the planner's budget constraint, λ, converts the fiscal revenue effect from dollars to welfare. λ measures the social value of a dollar in public funds. This value depends on what the planner uses the dollar for. As I formally show below, a natural lower bound for λ is the average marginal utility of consumption in the population. In this case, the planner would simply redistribute the dollar lump-sum across all individuals.

The optimal benefit function then ensures that there is no potential welfare improvement from any reform, i.e. that $\delta W(b; h) = 0$ for all deviations $h(\cdot)$. Starting from (6) the optimal benefit function therefore fulfills

$$\underbrace{\frac{\frac{1}{M(h)} \int \omega_i E_i \left[\sum_{t=0}^{T-1} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)} h(x^t)\right] di}{\lambda}}_{\text{Social value of the dollar}} = \underbrace{1 + \frac{B(h)}{M(h)}}_{\text{Multiplier}} \tag{9}$$

for all deviations $h(\cdot)$. A formal proof of this optimality condition can be found in Appendix A. Normalizing the welfare effect by the mechanical effect $M(h)$ provides the same metric for all reforms in equation (9), making an increase in the early retirement age comparable to a cut in pension levels. The thought experiment in equation (9) is therefore to take away one dollar from retirees through a specific reform. The LHS of (9) measures the social value of the dollar in the hands of the affected group. The RHS of (9) measures the multiplier of the reform, i.e. how much additional fiscal revenue is generated by behavioral responses of agents on top of the mechanically saved dollar. Formula (9) nests the Baily-Chetty formula for optimal unemployment insurance benefits as a special case as I illustrate in Appendix A.3.

Welfare Effect of a Reform. How is equation (9) useful for thinking about welfare effects of actual pension reforms? Suppose we have a specific reform that makes the pension system less generous, for instance a reduction in pension levels. Further suppose that we can determine the social value of the dollar (LHS of (9)) and the multiplier (RHS of (9)) of this reform. If we find that the multiplier is larger than the social value of the dollar, we can conclude that the reform is welfare improving.

$$\delta W(b;h) \gtreqless 0 \;\Leftrightarrow\; \underbrace{\frac{\frac{1}{M(h)}\int \omega_i E_i\left[\sum_{t=0}^{T-1}\frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)}h(x^t)\right]di}{\lambda}}_{\text{Social value of the dollar}} \lesseqgtr \underbrace{1+\frac{B(h)}{M(h)}}_{\text{Multiplier}}$$

If we find that the social value of the dollar is larger than the multiplier of the reform, we would have better not changed the system and the dollar is better left in the hands of the retirees. In that case the reform is welfare-reducing. If the pension system was optimal before the reform, the social value of the dollar and the multiplier would be equal.

But can we get credible estimates of the social value of the dollar and the multiplier of a reform? In principle we can estimate the multiplier of a reform with reduced-form methods. The social value of a dollar, however, is not directly related to observed moments in the data. In the end this is a judgment call of the planner, since it will always depend on the welfare weights a planner chooses. Nevertheless, knowing the multiplier of a reform is informative. The multiplier provides the benchmark of how high the social value of a dollar must be such that the reform is no longer welfare improving. For instance, assume reducing pension levels has a multiplier of 1.5, meaning that for each mechanical dollar there is an additional 50 cents reduction in expenditures due to behavioral response (e.g. individuals retire later and pay more taxes). Reducing pension levels is then welfare improving as long as the dollar in the hands of the affected retirees is valued at less than 1.5 dollars. To illustrate the magnitude of the social value of the dollar I also parameterize the model. Appendix A.5 provides the details for this exercise.

Lower Bound for the Social Value of the Dollar. A natural lower bound for the social value of the dollar is unity, meaning that one dollar in the hands of a retiree has a social value of at least one dollar. To derive this lower bound, we need to make two assumptions. First, assume that welfare cannot be improved through higher lump-sum taxation. This is a fairly mild assumption on the efficiency of the tax system and implies that λ is smaller than the average marginal utility of consumption in the population. Second, assume a utilitarian planner, i.e. $\omega_i = 1\,\forall i$.[9] With these two assumptions we can bound the social value of the dollar by the average marginal utility of consumption of individuals affected by the pension reform divided by the average marginal consumption in the population, i.e.

[9]This simplifies the argument and rules out that a planner would put zero weights on individuals, who are affected by the reform.

$$\text{Social Value of the Dollar} \geq \frac{\int E_i \left[\sum_{t=0}^{T-1} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)} h(x^t) \right] di}{\int E_i \left[\sum_{t=0}^{T-1} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)} \overline{h} \right] di} \geq 1 \tag{10}$$

where $\overline{h}$ is the lump-sum transfer, which would redistribute the mechanical effect from the reform $M(h)$ equally across all agents and periods. If we think that the marginal utility of consumption of retirees affected by the reform is larger than the average marginal utility of consumption in the population, then the social value of the dollar is larger than unity. Put differently, the social value of the dollar is larger than unity if we think that a transfer to retirees is socially more valuable than a transfer to all individuals. The formal argument for this lower bound is in Appendix A.2. In consequence, a reform with a multiplier below unity cannot be welfare-improving if we are willing to make these fairly mild assumptions.

Ranking Welfare Effects of two Reforms. Multipliers of reforms can be informative to rank the welfare effects of two reforms without exactly pinning down the social value of the dollar. For instance, if two reforms target the same individuals, the social value of the dollar is identical. Hence, the multipliers are sufficient to rank the two reforms and the reform with the higher multiplier is preferable. In practice, we do not have two reforms that target the exact same individuals. However, we might have two reforms where we can rank the social value of the dollar of the two reforms under mild assumptions on the planner's preferences. For instance, assume that we have two reforms A and B. Suppose that reform A reduces pension levels for high income individuals while reform B reduces pension levels for all individuals. In this case the social value of the dollar of reform A is lower than the social value of the dollar of reform B if we assume the planner has redistributive preferences (since it is less costly to take away one dollar from rich individuals). If reform A also has the higher multiplier we can conclude that reform A is preferable to reform B.

Formally, the difference of the welfare effects of the two reforms is the difference in the multipliers (first line in equation below) minus the difference in the social value of the dollar (second line). If the multiplier of reform A is larger than the multiplier of reform B, the first term is positive. If the planner has preferences for redistribution and reform A targets higher income individuals compared to reform B, then the social loss of reform B is larger and the term in the second line is positive. As a consequence, reform A is unambiguously preferable to reform B.

$$\begin{aligned}\delta W(b;h_A)-\delta W(b;h_B) &= \left(\left[1+\frac{B(h_A)}{M(h_A)}\right]-\left[1+\frac{B(h_B)}{M(h_B)}\right]\right)\\ &+ \frac{1}{\lambda}\left(\frac{\int \omega_i E_i\left[\sum_{t=0}^{T-1}\frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)}h_B(x^t)\right]di}{M(h_B)}-\frac{\int \omega_i E_i\left[\sum_{t=0}^{T-1}\frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)}h_A(x^t)\right]di}{M(h_A)}\right)\end{aligned}$$

Estimating the Multiplier. The ideal experiment to estimate the multiplier of a reform would be to observe a treatment and control group under the pre- and post-reform regime over their life-cycle. In this case the multiplier can be estimated in three steps:

1. Estimate Total Fiscal Effect $T(h)$ as the mean difference in life-cycle fiscal revenue between treatment and control group.
2. Calculate the Mechanical Fiscal Effect $M(h)$: Calculate life-cycle fiscal revenue in the control group under both the pre- and post-reform pension regime (holding behavior fixed). The difference in fiscal revenue is the mechanical effect.
3. Compute the Behavioral Fiscal Effect as residual $B(h) = T(h) - M(h)$.

Reduced-form methods, such as difference-in-difference or regression-discontinuity estimators, can deliver credible estimates on the total fiscal effect of a reform. The mechanical effect of a reform is straightforward to calculate. The challenge of estimating the multiplier, however, is the life-cycle aspect and the ideal experiment does not usually exist. Typically, individuals learn about changes in their pension rules around age 50 to 60 and not already at the beginning of their life. If they knew about the reform already at younger ages, forward-looking individuals might adjust their behavior in anticipation of the reform. Hence, the challenge is to estimate these potential anticipation effects of a pension reform. For some of the Austrian reforms, individuals learned about changes in their rules well in advance and I provide evidence that anticipation seems not to play a key role. In the empirical part of the paper, I therefore argue that we can get a good sense of the multiplier of a pension reform with reduced-form methods.

Extensions. In appendix A.4 I show that with multiple overlapping generations we want to estimate the exact same multiplier as in the one generation setup to evaluate welfare effects of pension reforms. The logic of the argument is that we want to design the optimal pension formula for each generation and hence need to know the multiplier of a reform within each generation.

The social value of the dollar now also accounts for the relative valuation of a dollar in the hands of the current generation versus in the hands of future generations. Appendix A.6 discusses how the welfare evaluation changes in case of general equilibrium effects, non-marginal changes or behavioral biases of agents. In presence of these effects, there is an additional term on the LHS of formula (9), i.e. the direct welfare effect looks different. The RHS is still the multiplier of the reform. In particular, non-marginal changes and behavioral biases imply that behavioral adjustments have first-order welfare effects. In case of non-marginal reforms we can sign the additional term on the LHS. Behavioral adjustments to non-marginal reforms are costly and increase the social value of the dollar. Hence, if a reform is welfare-reducing in the marginal change framework, the reform is even more welfare-reducing if we account for the direct costs of behavioral adjustments. With behavioral biases the sign of the additional term (the "bias correction term") depends on whether a reform reduces or amplifies the bias. For example, suppose individuals are myopic and a reform induces them to retire later and save more. In this case, the reform reduces the bias (myopic individuals save too little and retire too early from the viewpoint of a paternalistic planner) and hence accounting for this bias term reduces the social value of the dollar making the reform more attractive.

3 Institutional Background and Data

3.1 The Public Pension System in Austria

The Austrian public pension system covers all private-sector workers. The system is primarily financed as a pay-as-you-go system, but financial shortfalls have to be covered by the federal budget.[10] The other pillars of the pension system are of minor importance.[11] Public pensions are the main source of income for retirees and they replace on average 75% of the pre-retirement net earnings. Public pensions are calculated as a function of insurance years, experience, retirement age and past wages. Since 1985 Austria had seven major pension reforms that changed the benefit formula substantially. In general, the pension benefit formula consists of two parts, the assessment basis and the pension coefficient:

[10]In 2017, the share coming from the federal budget (the so-called "Bundesbeitrag") amounted to about 1.7% of GDP or 17.2% of overall spending for old-age pensions (HV, 2018).

[11]Funded company pension schemes, comparable to 401(k) plans in the U.S., are not mandatory and in 2007 less than 20% of the Austrian workforce was covered by such plans. Since 2002, there has been a new severance pay scheme where employers transfer 1.53% of the monthly salary to a pension account. The contribution rate to the pay-as-you-go pension system is 22.8%. Third-pillar pensions ("Prämienbegünstigte Zukunftsvorsorge") have only been available since 2003. For more details see Fink (2009).

$$\text{old-age pension} = \text{assessment basis} \times \text{pension coefficient}.$$

The assessment basis measures the average earnings over a specific period after applying a cap to earnings in each year. The assessment basis is comparable to the Average Indexed Monthly Earnings (AIME) in the U.S. system. The pension coefficient is the individual's replacement rate (in percent) that is applied to the assessment basis. The pension coefficient is a function of the number of insurance years and the claiming age.[12]

Individuals with more than fifteen insurance years are eligible to claim old-age pensions. The normal retirement age (NRA) is 60 for women and 65 for men. The early retirement age (ERA) was 55 for women and 60 for men and was step-wise increased to 60 for women and to 65 for men. Since 2005, individuals with more than 37.5 insurance years can still retire early at age 62. In Austria, pension eligibility and payments always depend on individual accounts. It is not possible to qualify for retirement benefits through one's spouse's contribution record.

3.2 Pension Reforms since 1985

Since 1985 Austria has had seven major pension reforms. The seven reforms changed all relevant margins of the pension benefit formula at least once and Austrian policy makers phased in most policy changes yielding vast quasi-experimental variation. The early reforms in 1985 and 1988 reduced pension levels by adjusting the definition of the assessment basis. The pension reforms in 1993, 1996, 2000 and 2003 changed the actuarial fairness by adjusting the pension coefficient to penalize early claiming. The reforms in 2000 and 2003 also increased the early retirement age. Table 1 provides an overview of the seven pension reforms. The changes in the assessment basis, pension coefficient and retirement ages were all phased-in with complicated transition rules. This creates quasi-experimental variation in pension rules, which I exploit in my empirical analysis. I explain the complicated transition rules and how I exploit them in section 4 in detail. In the following, I shortly describe how the three main determinants of the pension formula, the assessment basis, the pension coefficient and the retirement age, changed over time.

Assessment Basis. Before 1985, the assessment basis was calculated as the average earnings in the last five years. The pension reform in 1985 increased the assessment period from the last five to the last ten years. In 1988, this period was further increased from the last ten years to

[12]Insurance years include both contribution years (i.e., periods of employment, including sick leave) and non-contributory periods of labor force participation (e.g., unemployment).

the last fifteen years. These reforms reduced the pension level for most individuals, since the average earnings of the last fifteen years are lower than the average earnings in the last five years for individuals with increasing wage profiles due to seniority effects. In 1993, the assessment basis was again changed from the average earnings of the last fifteen years to the average earnings of the best fifteen years. For most individuals, the best fifteen years are exactly the last fifteen years and this reform did not substantially change pension levels. Starting in 2004, the assessment period was step-wise increased from the best fifteen years to the best 40 years. This change was phased in between 2004 and 2028; each year the assessment period increased by one year.

Pension Coefficient. Before 1993, the pension coefficient only depended on the number of insurance years and was independent of the claiming age. Up to 30 insurance years, each insurance year increased the pension coefficient by 1.9 percentage points. Above 30 insurance years, the pension coefficient increased by 1.5 percentage points with each additional insurance year up to a maximum of 80%. The maximal pension coefficient was therefore reached at 45 insurance years. The pension reform in 1993 introduced a bonus for claiming pensions after the early retirement age. This made pensions more generous for individuals who claim after the early retirement age. For each month after the early retirement age the pension coefficient was scaled up by a certain factor. For example, claiming five years after the early retirement age scaled the pension coefficient up by a factor of 1.11. The reform in 1996 introduced a penalty for claiming before age 56 for women and 61 for men. This penalty for claiming early depends on the number of insurance years. The pension reforms in 2000 and 2003 further changed the penalty. Since 2000, each insurance year increases the pension coefficient by two percentage points and the penalty for each year claiming before the NRA is set at three percentage points with a maximum penalty of ten percentage points. The reform in 2003 further increased this penalty to 4.2 percentage points per year with a maximum of 15%. These changes were phased in over time with complicated caps on benefit losses with respect to prior rules.

Retirement Age. The reforms in 2000 and 2003 also increased the early retirement age in steps for men from 60 to 65 and for women from 55 to 60. The normal retirement age will be increased from 60 to 65 for women in 2024. Interestingly, this change was already enacted in 1993. While the actual change has not taken place yet, the announcement almost 30 years ago allows studying anticipation effects over long periods of time and shedding light on forward-looking behavior of individuals with respect to increases in the normal retirement age.

Table 1: Austrian Pension Reforms

Reform	Assessment Basis	Pension Coefficient (%)	Retirement Age
Rules before 1985	indexed earnings of last 5 years	only depends on insurance years (IY) $\mathrm{PC} = \begin{cases} 50 & \text{if } 15 \leq \mathrm{IY} < 25 \\ 51 + 1.2(\mathrm{IY} - 25) & \text{if } 25 \leq \mathrm{IY} < 30 \\ 57 + 1.5(\mathrm{IY} - 30) & \text{if} \\ 79.5 & \text{if } 45 \leq \mathrm{IY} \end{cases}$	ERA: 55 women, 60 men NRA: 60 women, 65 men
1985	indexed earnings of last 10 years	reduction in pension coefficient for individuals with less than 25 insurance years ($1.9 * (\mathrm{IY})$ instead of the flat 50%).	-
1988	indexed earnings of last 15 years	-	-
1993	indexed earnings of best 15 years	pension coefficient depends on claiming age. Bonus for claiming after ERA, PC is scaled up by a factor for each month of claiming after ERA.	Increase of NRA for women from 60 to 65, but implementation starts in 2024.
1996	-	Penalty for claiming before age 56/61. Penalty depends on number of insurance years and only applies to individuals with $33 \leq \mathrm{IY} < 40$.	-
2000	-	increase of penalty for claiming before NRA (from 2 to 3 pp for each year before NRA with caps on loss)	increase of ERA from 55 to 56.5 for women, from 60 to 61.5 for men
2003	indexed earnings of best 40 years (increase by 1 year every year starting in 2004)	reduction in PC from 2 pp for each insurance year to 1.78 pp. increase in penalty for claiming early from 3 pp to 4.2 pp for each year before NRA.	increase of ERA from 56.5 to 60 for women, from 61.5 to 65 for men
2004	introduction of individual pension accounts based on life-time earnings		Corridor pension: early retirement at age 62 again possible with more than 37.5 insurance years.

3.3 Other Social Insurance Programs

Apart from old-age pensions, there are three other important social insurance programs in Austria: disability insurance (DI), sickness insurance (SI), and unemployment insurance (UI). The DI program provides partial earnings replacement to workers below the full retirement age who have accumulated at least 5 insurance years within the last 10 years and have a health impairment that is considered severe. DI benefits are calculated in a similar fashion as old-age pensions (based on the assessment basis and the pension coefficient) and replace approximately 70 percent of pre-disability net earnings up to a maximum of around €4,500 per month. SI benefits cover workers with temporary illness, which last longer than 12 weeks. SI benefits replace approximately 65% of the last net wage and the benefit duration is 52 weeks for individuals who have worked at least 6 months in the previous 12 months, and 26 weeks otherwise. UI benefits replace approximately 55% of the wage on the last job subject to a minimum and maximum. Unemployed below age 50 receive at most 39 weeks of regular UI benefits, job losers above age 50 can claim benefits for up to 52 weeks provided they have paid UI contributions for at least 9 years in the last 15 years.[13]

3.4 Austrian Social Security Data (ASSD)

In my empirical analysis, I use the Austrian Social Security Data (ASSD), which is described in Zweimüller et al. (2009). The ASSD cover the universe of private-sector employees since 1972 and contain detailed information on labor market status, earnings and demographic variables. This data set allows constructing social security benefit receipt. Based on individual's earnings records and labor market histories, I calculate old-age pensions (OA), disability pensions (DI), unemployment benefits (UI) and sick leave benefits (SI) as well as social security contributions.[14] I define fiscal revenue as payroll taxes minus old-age (OA) pensions, disability insurance (DI) benefits, unemployment insurance (UI) benefits, and sick-leave insurance (SI) benefits

[13] After UI benefit exhaustion individuals can apply for "unemployment assistance" (UA), which is means-tested. UA benefits last for an indefinite period and replace around 70% of regular UI benefits. However, I do not observe UA take-up in the data.

[14] I do not directly observe pension and benefit payments but the corresponding labor market status and then construct pension and benefit payments based on my calculated benefits. I observe actual OA and DI pension payments for individuals, who receive OA or DI pensions in 2001 or start claiming after 2001. To verify my pension calculations, I compare my calculated OA and DI pensions with the actual payments for this subsample. Figures 42 to 47 in Appendix E show that on average my calculated pensions track actual pensions and that there is no systematic error in my calculations across pension levels or over time.

$$\text{Fiscal Revenue} = \text{Pay roll taxes} - (\text{OA pension+DI benefits+ UI benefits+SI benefits}) .$$

This definition of fiscal revenue treats social security as a closed system and ignores the potential effects of a pension reform on other transfer programs and other tax revenue. Ideally, I would want to include all taxes and transfers that affect the government budget constraint. However, I am limited by the data. Some taxes, like VAT, and some transfers, like social assistance, simply cannot be observed in the data. Moreover, I cannot reliably calculate the additional revenue from income taxation because labor income in the ASSD is top and bottom coded.[15] By ignoring other forms of taxation, I tend to underestimate the multipliers.

4 Empirical Evidence

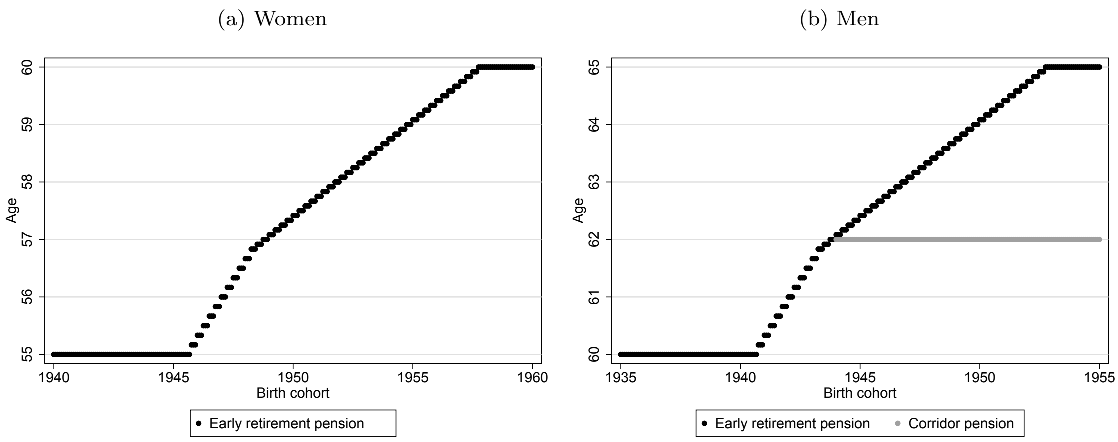
4.1 Early Retirement Age

Policy Variation. The pension reforms in 2000 and 2003 increased the early retirement age (ERA) step-wise from 60 to 65 for men and from 55 to 60 for women. Figure 2 plots the variation in ERA by date of birth. The 2000 pension reform increased the ERA by 1.5 years and the increase was phased in by quarter of birth. For men born between October 1940 and September 1942, the ERA increased by two months every quarter of birth from 60 to 61.5. For women born between October 1945 and September 1947, the ERA increased by two months every quarter of birth from 55 to 56.5. Men with at least 45 contribution years and women with at least 40 contribution years were unaffected by this reform and could still retire at age 60 and 55 respectively. The 2000 pension reform was debated in June 2000 and put into practice in October 2000. The 2003 pension reform further increased the ERA from 61.5 to 65 for men and from 56.5 to 60 for women and the increase was again phased-in by quarter of birth. First, the ERA increased by two months for each quarter of birth for men born between January and June 1943 and for women born between January and June 1948. Then, the ERA increased by one month for each quarter of birth for men born between July 1943 and December 1952 and women born between July 1948 and December 1957. The 2003 pension reform was in parliament in June 2003 and became effective on January 2004.

The pension reforms in 2000 and 2003 also changed other margins of the pension formula.

[15] Pay roll taxes only apply to the uncensored part of the income distribution and I can therefore calculate them.

Figure 2: Early Retirement Age Variation by Quarter of Birth



Notes: This figure shows the variation in the early retirement age by date of birth. The 2000 pension reform increased the ERA for women from 55 to 56.5 and for men from 60 to 61.5 stepwise by 2 months for each quarter of birth. The 2003 pension reform further increased the ERA to 60 for women and 65 for men. The 2004 pension reform reintroduced the possibility of early retirement at age 62 for individuals with more than 37.5 insurance years ("corridor pension").

However, for the 2000 pension reform these other changes are of minor importance for my empirical strategy. Importantly, for the 2000 pension reform the pension formula did not change if individuals claimed at their ERA. The 2003 pension reform is more challenging in this respect. In appendix B, I discuss these other changes in detail.

In the main text I focus on the 2000 pension reform, which increased the ERA from 55 to 56.5 for women and from 60 to 61.5 for men. The analysis of the 2003 pension reform is in Appendix C.

Sample Selection. For the 2000 pension reform, my main sample consists of all women born between 1945 and 1947 and all men born between 1940 and 1942. For the 2003 pension reform, my main sample consists of all women born between 1948 and 1957 and all men born between 1943 and 1952. In all samples, I exclude individuals who have worked in publicly-owned industries (public administration, public transportation, and education), as public sector workers are covered by a separate pension system with different eligibility rules. I further exclude self-employed individuals and individuals who have spent any time working in jobs defined as heavy labor, as they might be eligible for a special heavy labor pension. Furthermore, I exclude women with more than 40 contribution years and men with more than 45 contribution years, since they are exempt from the increase in the ERA.

4.1.1 Descriptive Evidence

Figure 3 plots the percent of women in retirement, employment, unemployment and the residual category by age for different birth cohorts. The vertical lines indicate the cohort specific ERA. Panel (a) shows that around 10 percent of women are on disability benefits before age 55. Right at the ERA retirement take-up jumps up and around 50 percent of all women are in retirement. Retirement then gradually increases with age and at the NRA almost all women are retired. Panel (a) also indicates that the increase in the ERA simply shifts this retirement profile to the right.[16] Panel (b) shows that employment rates significantly drop at the ERA and the increase in ERA shifts the employment profiles to the right. Panel (c) illustrates that the unemployment rate increases to almost 20 percent in the year before the ERA and then drastically drops at the ERA. Again the increase in the ERA seems to shift this profile in parallel to the right. Panel (d) suggests that the residual category does not systematically vary across birth cohorts with different ERA.

Figure 4 displays the labor market - age profiles of men. Panel (a) shows that labor force participation of older men in Austria is very low. At age 59, already 50 percent of all men retired through disability pensions and only 30 percent are in employment. This is due to the generous disability insurance system with relaxed eligibility rules for individuals older than 57.[17] At the ERA around 90 percent are in retirement and less than 10 percent are still working. As for women, the increase in the ERA seems to shift all profiles to the right.

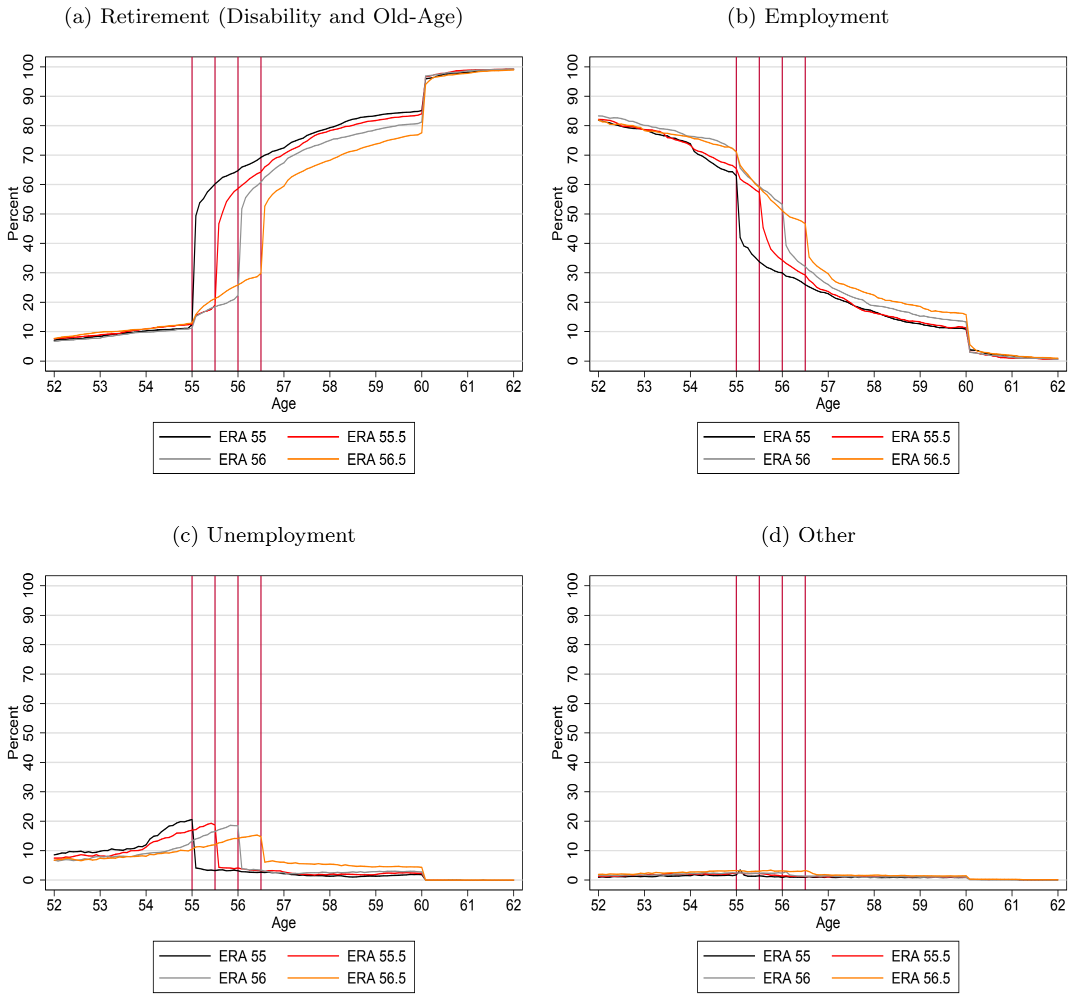
4.1.2 Empirical Strategy

I exploit the variation in the ERA by quarter of birth in a cohort difference-in-difference specification. This approach compares younger and older cohorts, who face different ERA rules, over time. My control group consists of individuals born in the last quarter, which is not affected by the reform. Individuals in the control group can still retire at the pre-reform ERA. For the 2000 pension reform I have nine different treatment groups. Each quarter of birth with a different

[16]However, the figure also shows that retirement take-up increases already before the ERA. Only around a third of this increase is explained by disability insurance take-up. Hence, there are women who claim retirement prior to the new ERA. This is due to measurement error in the number of contribution years, which leads to some misclassification in eligibility rules. The number of contribution years is not directly observable and I calculate it based on labor market histories. The calculation of the contribution years is not exact because some labor market histories are censored in 1972 and I impute contribution years following the approach in Staubli and Zweimüller (2013). Moreover, the definition of contribution years is not straightforward. For instance, individuals can buy in contribution years for some of their education and I might not observe all of this. As a consequence, there are some women (men) who have actually more than 40 (45) contribution years, but I calculate less than 40 (45) contribution years for them. This problem seems to be more severe for women than for men. However, this is not a problem for my identification strategy. The misclassification reduces the absolute magnitudes of my estimates, since there is not 100 percent compliance. However, I am primarily interested in the relative size of effects (behavioral vs. mechanical fiscal effect) and the relative size is not affected by this misclassification.

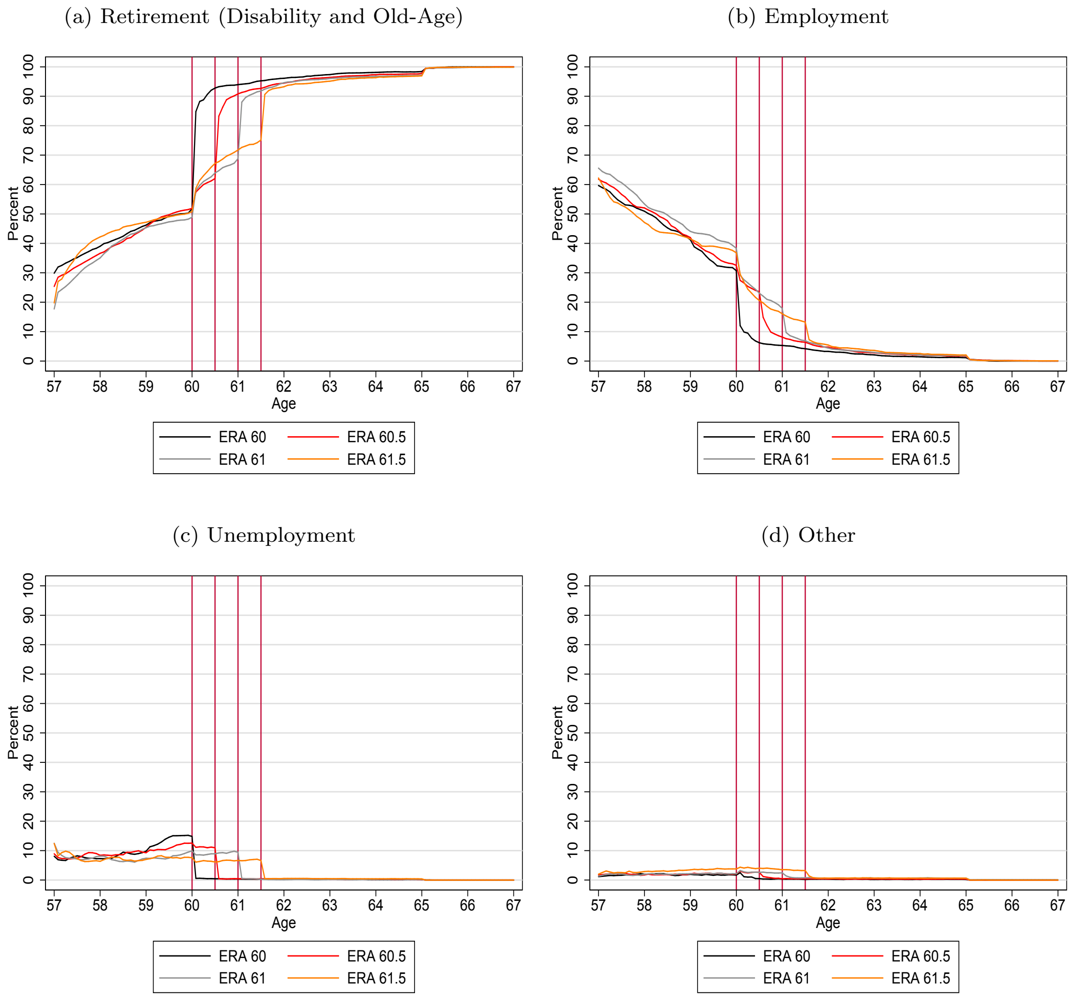
[17]For a detailed discussion of the Austrian DI system see Staubli (2011); Haller et al. (2019)

Figure 3: Women's Labor Market Status by Age



Notes: This figure shows the labor market profiles for women born in different quarters. The vertical lines indicate the cohort specific ERA.

Figure 4: Men's Labor Market Status by Age



Notes: This figure shows the labor market profiles for men born in different quarters. The vertical lines indicate the cohort specific ERA.

ERA forms a separate treatment group. I then compare the outcomes of the treatment groups and the control group at all ages between three years before the ERA and three years after. The comparison between the control group and the first treatment group identifies the effect of increasing the ERA by two months. The comparison of the second treatment group with the control group identifies the effect of increasing the ERA by four months and so on up to the ninth treatment group, where we identify the effect of increasing the ERA by 1.5 years.[18] I implement this comparison by estimating the following regression

$$Y_{it} = \alpha + \sum_{j=1}^{9} \sum_{k=ERA-36}^{ERA+36} \beta_{kj} * Treat_{ij} * I[age_{it} = k] + \sum_{k=ERA-36}^{ERA+36} \kappa_k * I[age_{it} = k] + \sum_{j=1}^{9} Treat_{ij} + \lambda_t + \varepsilon_{it} \tag{11}$$

where i denotes individuals and t year-months. Y_{it} is the outcome variable of interest (such as net transfer payments, and labor supply measures like indicators for working or retirement). $I[age_{it} = k]$ are dummies for age at a monthly frequency and control for age-specific levels in the outcome variable. $Treat_{ij}$ are dummies, which indicate to which of the nine treatment groups the individual belongs. This is determined by their quarter of birth. λ_t are time dummies at monthly frequency to capture common time shocks and seasonal effects. β_{kj} then identifies the effect at age k in treatment group j.

I am interested in three main outcomes. First, how does the policy change affect fiscal revenue? For this I construct for each individual the net transfer as payroll taxes minus the sum of retirement, unemployment, disability and sick leave benefits as described in section 3.4. Second, I decompose the fiscal revenue effect into the behavioral and mechanical fiscal effect to get the multiplier of the reform. Third, I am interested in understanding the behavioral fiscal effect, i.e. how individuals adjust their labor market decisions in response to the change in the ERA. I construct labor market status indicators for employment, retirement, unemployment, disability pension and sick leave and also look at transitions.

4.1.3 Results

Fiscal Revenue Effect. Figure 5 plots the β_{kj}-coefficient estimates from regression (11) for the total fiscal revenue (measured in Euros) for women. Panel (a) illustrates the effect of increasing the ERA by 2 months from 55 to 55 and 2 months. Panel (b) presents the estimates for increasing

[18] As a robustness check I compare only adjacent quarters of birth and run 9 separate difference-in-difference regressions. The counterfactual is then to increase the ERA by two months but starting at different ages. The results of this exercise are in Appendix B.2.

the ERA by 6 months, in panel (c) the ERA is increased by one year and in panel (d) by 1.5 years. The red line indicates the pre-reform ERA at age 55, the grey solid line shows the new ERA and the dashed grey line is located at the age individuals learned about the reform. There are three main takeaways from these figures. First, increasing the ERA has a significant and positive effect on fiscal revenue. For each month of increasing the ERA the net fiscal revenue per capita increases by around 650 Euros per month. Second, there are no effects before individuals knew about the reform (at ages to the left of the dashed line), which implies that trends of the control and treatment groups are parallel pre-reform. Third, the effects are limited to the age window where the ERA increased. There are no anticipation effects before age 55. Some individuals knew more than two years in advance that their ERA is increased by more than one year (panel (d)). One could expect that individuals adjust to this increase already before age 55. For instance, with a higher ERA unemployed individuals have a stronger incentive to search for a new job, since they might run out of UI benefits before they reach the ERA. However, there is no strong evidence for such anticipation effects. Moreover, the effects vanish exactly at the new ERA and hence there are no long lasting effects beyond the new ERA.

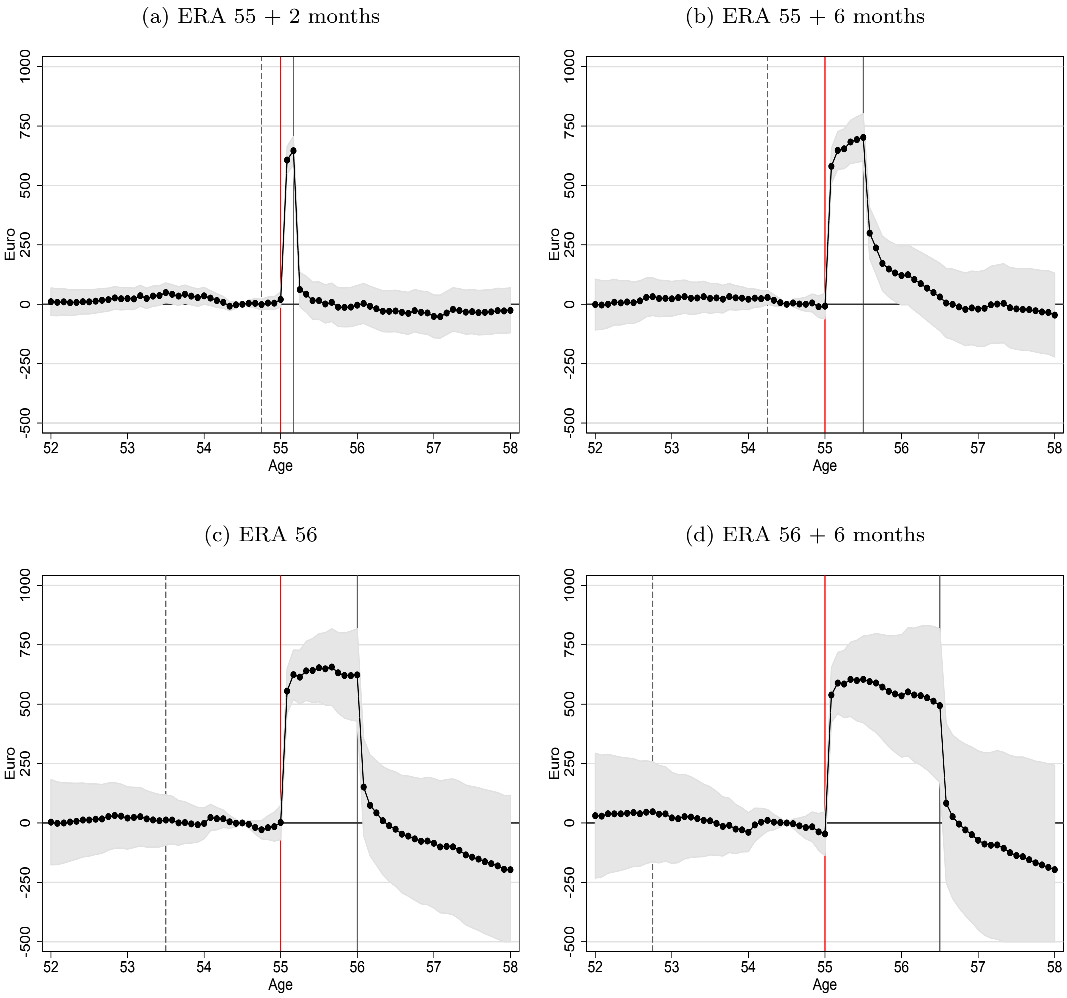
Figure 6 shows the fiscal revenue effects for men. The patterns are identical to the women's patterns. Only the magnitude of the effects is slightly higher, since men tend to have higher pensions. The patterns for the other five treatment groups also look very similar, these figures can be found in Appendix B.1.

Multiplier. The crucial question for welfare analysis is how much of the fiscal revenue effect is purely mechanical and how much is due to behavioral adjustments. The mechanical fiscal effect is simply the reduction in old-age pension payments between the old and new ERA. The counterfactual for the mechanical fiscal effect is that individuals would behave the same as before the reform. That is, they would stop working at the old ERA and would not receive their old-age pension until they reach the new ERA. However, individuals respond to the change in the ERA by working longer or substituting to other benefits. The behavioral fiscal effect therefore consists of the additional pay roll tax revenue minus the additional expenditures in UI, DI and SI benefits.[19]

[19] I construct the mechanical fiscal effect by calculating the old-age pension expenditures in the control group between the old and new ERA and subtract the old-age pension expenditures in the treatment group in that age window. I need to subtract the old-age pension expenditures in the treatment group because some individuals in my treatment groups can still retire early if they have more than 40/45 contribution years. Without this correction I would overestimate the mechanical effect.

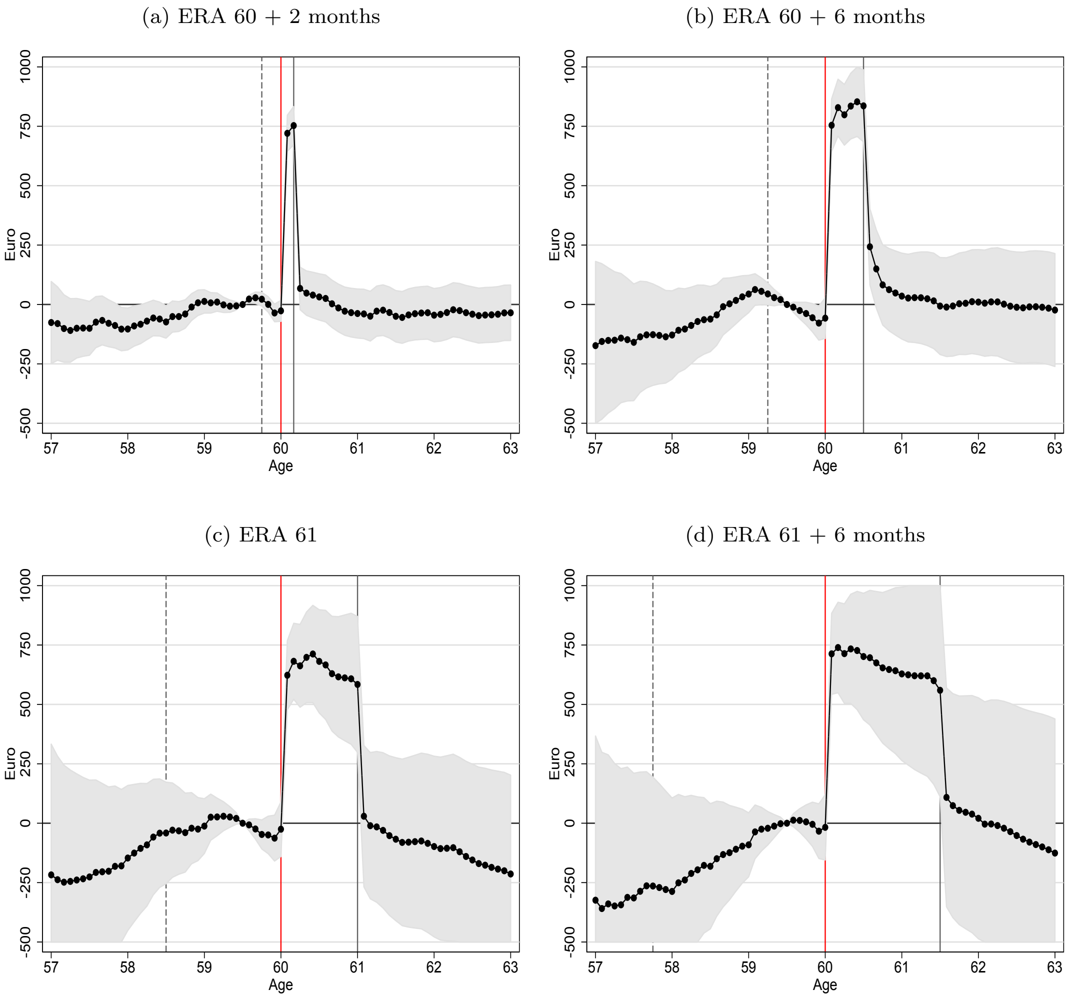
The alternative way to calculate the mechanical fiscal effect is to directly sum up the difference-in-difference estimates of the old-age pension expenditures between the old and the new ERA. The difference between these two approaches are minimal, the mechanical fiscal effects only differ by a few Euros.

Figure 5: DiD Estimates Fiscal Revenue by Age for Women Reform 2000



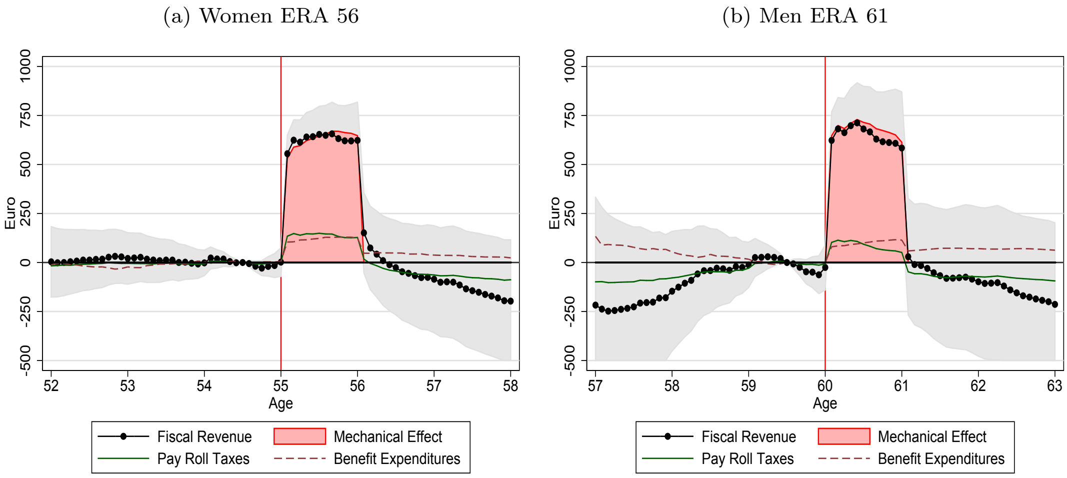
Notes: This figure plots the β_{kj}-coefficients from regression (11) for the total fiscal revenue (measured in Euros). Panel (a) shows the effect of increasing the ERA by 2 months, Panel (b) for a 6 month increase, Panel (c) for a one year increase and Panel (d) for a 1.5 year increase. The red line indicates the pre-reform ERA, the grey solid line shows the new ERA. The dashed grey line is located at the age individuals learned about the reform. Hence, effects between the dashed grey line and the red line could be interpreted as anticipation effects of the reform.

Figure 6: DiD Estimates Fiscal Revenue by Age for Men Reform 2000



Notes: This figure plots the β_{kj}-coefficients from regression (11) for the total fiscal revenue (measured in Euros). Panel (a) shows the effect of increasing the ERA by 2 months, Panel (b) for a 6 month increase, Panel (c) for a one year increase and Panel (d) for a 1.5 year increase. The red line indicates the pre-reform ERA, the grey solid line shows the new ERA. The dashed grey line is located at the age individuals learned about the reform. Hence, effects between the dashed grey line and the red line could be interpreted as anticipation effects of the reform.

Figure 7: Mechanical Effect ERA Reform 2000



Notes: This figure plots the estimated total fiscal revenue effect (black dots) and the mechanical fiscal effect (red area). For both women and men the fiscal revenue effect is purely mechanical. The green line plots the additional pay roll tax revenue and the maroon dashed line plots the additional benefit expenditures. The additional revenue from individuals working longer is offset by the additional expenditures from individuals substituting to UI, SI and DI benefits. These two effects cancel, leading to no additional savings from behavioral adjustments of individuals.

Figure 7 plots the estimates for the total fiscal effect (black dots) and the mechanical effect (red area) for women with ERA 56 and men with ERA 61. The figure shows that the large and positive fiscal revenue effect of increasing the ERA is purely mechanical, implying that the behavioral fiscal effect is zero. Why are there no additional savings from individuals working longer and paying additional taxes? There is a positive fiscal effect through additional pay roll tax revenue (green line in the figure), but there are also additional expenditures from individuals substituting to UI, DI and SI benefits (red dashed line in the figure). These two effects cancel and the net fiscal effect from behavioral adjustments is zero.

Figures 20 to 23 in Appendix B.1 show the decomposition of the total fiscal effect for the other treatment groups. The takeaway from these figures is that the total fiscal revenue effect is in all groups mostly driven by mechanical effect. Interestingly, at higher ERAs the negative fiscal effect of additional benefit payments starts to dominate the positive pay roll tax effect, which leads to a negative behavioral fiscal effect and a multiplier below one. Table 2 shows the total fiscal effect and the total effect for each treatment group. The column "Fiscal Revenue Effect" in table 2 is the sum of the significant β_{kj}-estimates from regression (11) for fiscal revenue. I abstract from discounting here because all effects are within a narrow time window of 1.5 years. The column "Mechanical" is the sum of the mechanical fiscal effect between the old and new ERA (red area in figure 7).

The behavioral fiscal effect is then calculated as the difference between the fiscal revenue effect and the mechanical effect. Table 2 reveals increasing the ERA is an effective policy to increase fiscal revenue. A two months increase in the ERA generates an increase in net fiscal revenue of 1200 - 1400 Euros per capita. However, this total fiscal effect is purely mechanical (column two). In consequence the behavioral effect is small or even negative. This leads to multipliers that are centered around one and in some cases as low as 0.9. This means for taking one dollar away by increasing the ERA, fiscal revenue increases by around one dollar or in the worst case by only 90 cents.

The multipliers for women are decreasing with the increase in the ERA. One might worry that this is primarily driven by time trends. However, I also find multipliers around one in the robustness check in Appendix B.2, where I only compare adjacent quarters of birth.

Table 2: Multipliers for the ERA Reform in 2000

ERA	Fiscal Revenue Effect	Mechanical	Behavioral	Multiplier (1+B/M)
Women				
ERA 55 + 2 months	1253	1120	133	1.12
ERA 55 + 4 months	3890	2474	1417	1.57
ERA 55 + 6 months	4949	3763	1185	1.32
ERA 55 + 8 months	5245	4970	275	1.06
ERA 55 + 10 months	6446	6230	216	1.03
ERA 56	7530	7578	-49	0.99
ERA 56 + 2 months	8337	8565	-228	0.97
ERA 56 + 4 months	8699	9694	-995	0.90
ERA 56 + 6 months	10069	11143	-1074	0.90
Men				
ERA 60 + 2 months	1473	1448	25	1.02
ERA 60 + 4 months	2552	3059	-507	0.83
ERA 60 + 6 months	5070	4855	216	1.04
ERA 60 + 8 months	6584	6433	152	1.02
ERA 60 + 10 months	7728	7611	117	1.02
ERA 61	7774	8173	-399	0.95
ERA 61 + 2 months	11102	10314	788	1.08
ERA 61 + 4 months	11576	10809	768	1.07
ERA 61 + 6 months	11918	11448	470	1.04

Labor Market Responses. With a higher ERA, unemployed individuals have a stronger incentive to search for a job and find employment, since they might run out of UI benefits before they reach the ERA. However, this does not happen. Individuals respond to an increase in the ERA in a very simple way. Individuals, who are employed before the old ERA, remain employed until they reach their new ERA or lose their job. Individuals, who are unemployed before the old ERA, remain unemployed until they reach their new ERA. I do not find any evidence for an increased transition from unemployment to employment at least in the short run. Figure 8 shows the levels for employment, unemployment and retirement. The 40 percentage points reduction in retirement is accompanied by a 20 percentage points increase in employment and a 20 percentage points increase in unemployment. Figure 9 plots the transitions and shows that this is driven by individuals keeping their jobs or remaining unemployed. There is no increased transition from unemployment to employment or the other way around.[20]

4.1.4 Welfare Implications

The analysis showed that increasing the ERA has a multiplier of around one. Hence, increasing the ERA is welfare improving only if the social value of the dollar in the hands of an early retiree is below one. The theoretical discussion in section 2 showed that a natural lower bound for the social value of the dollar is one. Hence, it is unlikely that increasing the ERA is welfare-enhancing.

It is still important to understand the distributional effects of increasing the ERA, i.e. to understand who retires early. Figure 10 plots the distribution of average income in the last 15 years (grey bars) and the share of individuals in that income bin, who retire early (red bars). The figure reveals that early retirement is prevalent across the distribution. An ERA reform therefore affects individuals from everywhere in the income distribution. This makes it harder to argue that the social value of the dollar is below one. Based on the argument in section 2, the average marginal utility of consumption of retirees would need to be lower than the average marginal utility of consumption in the population for the ERA reform to be welfare improving.

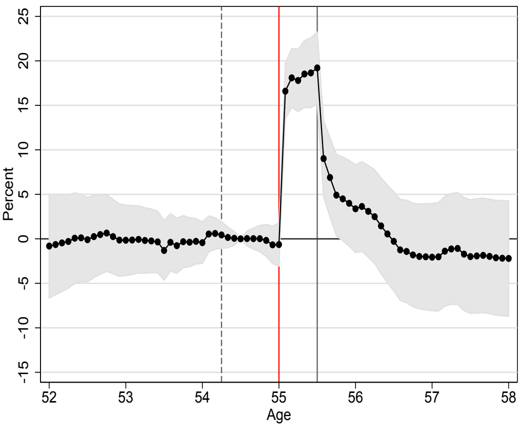
4.2 Pension Levels

An individual's pension is determined by her assessment basis multiplied with her pension coefficient. The assessment basis measures the average earnings over a specific period after applying a cap to earnings in each year. The pension coefficient is the individual's replacement rate (in percent) that is applied to the assessment basis. Before 1985 the assessment basis was the average earnings in the last 5 years. The pension reforms in 1985 and 1988 increased the assessment period step-wise from the last 5 years to the last 15 years. This makes old-age pensions on average less

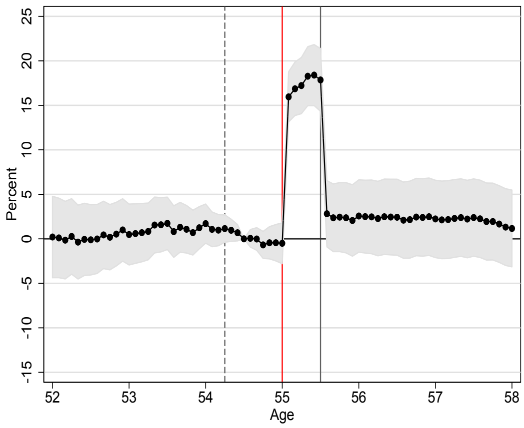
[20] In the longer run, there might be a positive employment effect before the old ERA. Figures 24 to 28 in appendix B show that for ERA 56 (60) and higher there is a downward sloping trend in unemployment before the ERA. It is, however, hard to tell whether this is a time trend or an effect of the reform.

Figure 8: DiD Estimates Labor Market Status by Age: Women ERA 55 + 6 months

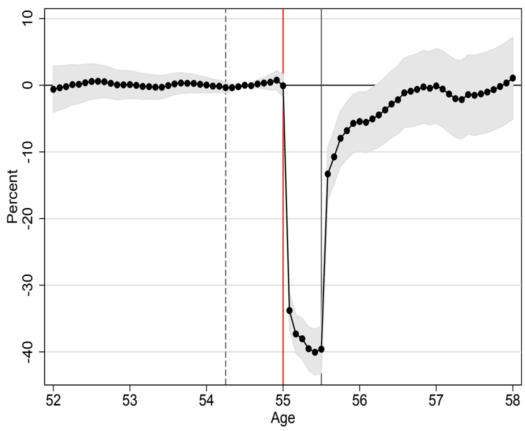
(a) Employment



(b) Unemployment

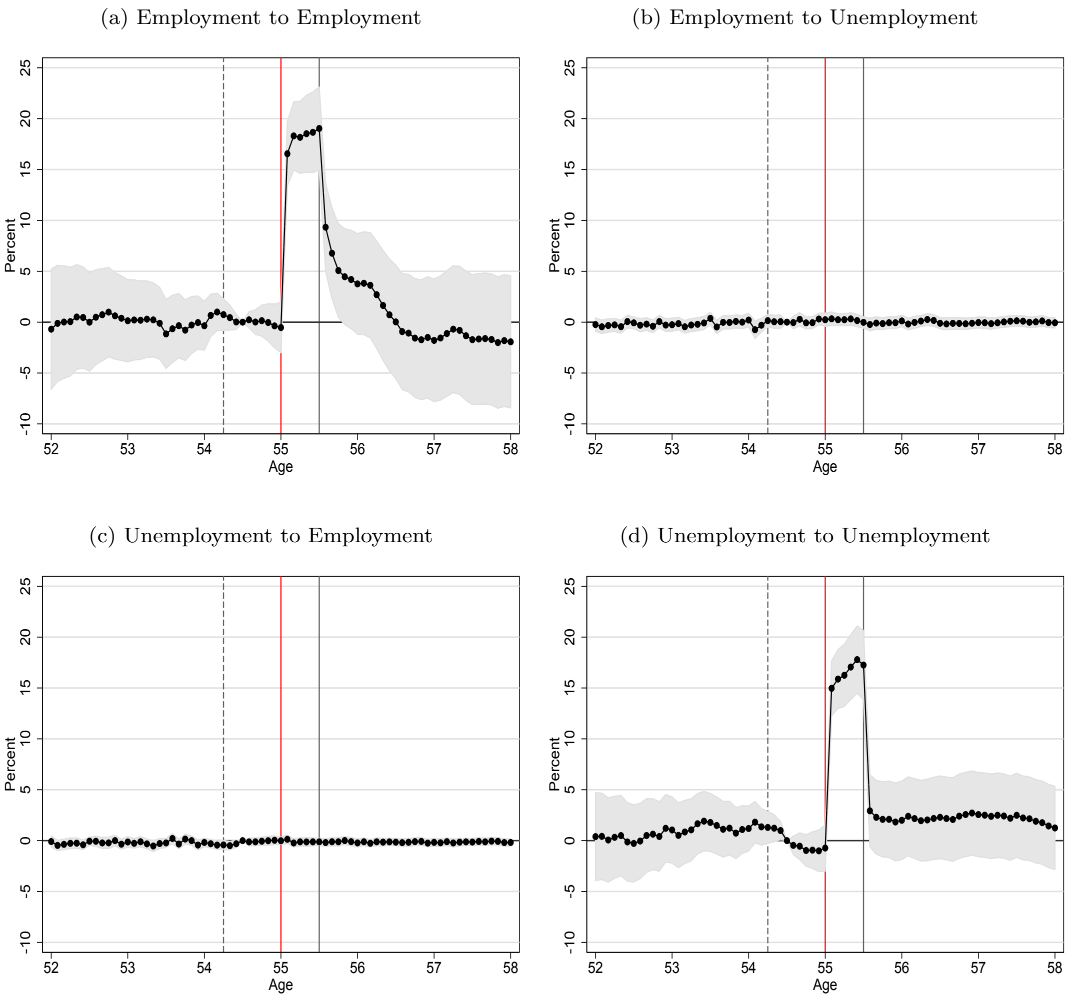


(c) Retirement (OA and DI)



Notes: This figure shows the DiD estimates for employment, unemployment and retirement. The 40 percentage points reduction in retirement is accompanied by a 20 percentage points increase in employment and a 20 percentage points increase in unemployment.

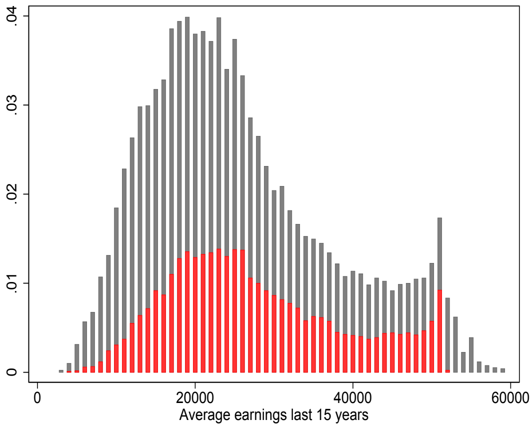
Figure 9: DiD Estimates Labor Market Transitions: Women ERA 55 + 6 months



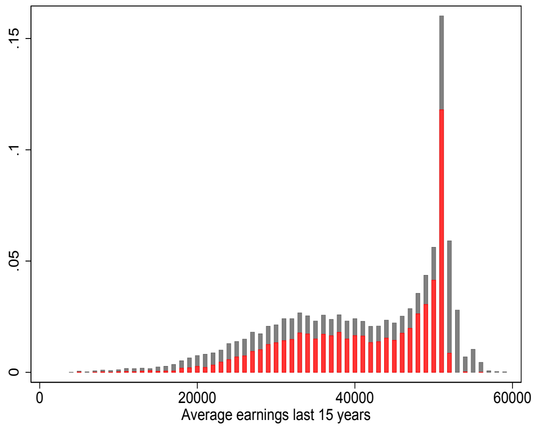
Notes: This figure shows the DiD estimates for labor market transitions.

Figure 10: Early Retirement across the Income Distribution

(a) Women



(b) Men



Notes: This figure plots the distribution of average income in the last 15 years (grey bars) and the share of individuals in that income bin, who retire early (red bars).

generous because of seniority wage profiles, however, it does not lead to a uniform cut in pension levels. Individuals across the income distribution are affected differentially. I focus on the change of the assessment period from the last 10 to the last 11 years implemented in the 1988 pension reform. This change provides a clean design as I outline in the next paragraph.

Policy Variation. In 1985, the assessment period was changed from the last 5 years to the last 10 years. This change was phased in over time. Between January and April 1985 the assessment period was either the last 5 or last 7 years, whatever was more favorable. Between May and December 1985, the assessment period was the last 7 years. In 1986, the assessment period was the last 9 years and in 1987 this was extended to the last 10 years. The 1985 pension reform also substantially reduced the pension coefficient for individuals with less than 30 insurance years. [21]

The 1988 pension reform changed the assessment period from the last 10 years to the last 15 years. This change was phased in by birth cohorts and over time. Table 3 describes the exact transition rules. The assessment period remains at 10 years for men born before 1.1.1928 and women born before 1.1.1933. Individuals born after these dates face an assessment period of 11 years in 1988. For men born in 1928 and women born in 1933, the assessment period remains at 11 years in the following years (unless the longer assessment period is more favorable). For

[21] Pre-reform individuals with less than 30 insurance years have a flat 50% pension coefficient. After the reform, the pension coefficient is calculated as 1.9 times the number of insurance years. Hence, an individual with 15 insurance years has a pension coefficient of 50% pre-reform and after the reform her pension coefficient is 28.5%.

the other cohorts, the assessment period increases each year by one year until they reach age 60. Only for the male cohort of 1928 and the female cohort of 1933, the rules changed once and then remained in place for the following years. I focus my analysis on these two cohorts and the 1988 pension reform, since this is the cleanest design with a clear counterfactual. [22]

The 1988 pension reform was announced and implemented in a very short time window. In July 1987, it was announced that a new pension reform will be debated and that the reform should be implemented in 1989. However, in October 1987 the federal ministry of labor, social affairs, health and consumer protection sent out a reform proposal. The proposal passed legislation in November. The transition rules for the increase in the assessment period was not in the initial proposal and was added during the legislative process. The new rules were in place on 1.1.1988. Hence, men born in January 1928 and women born in January 1933 only learned two months in advance that they will be treated by the reform.

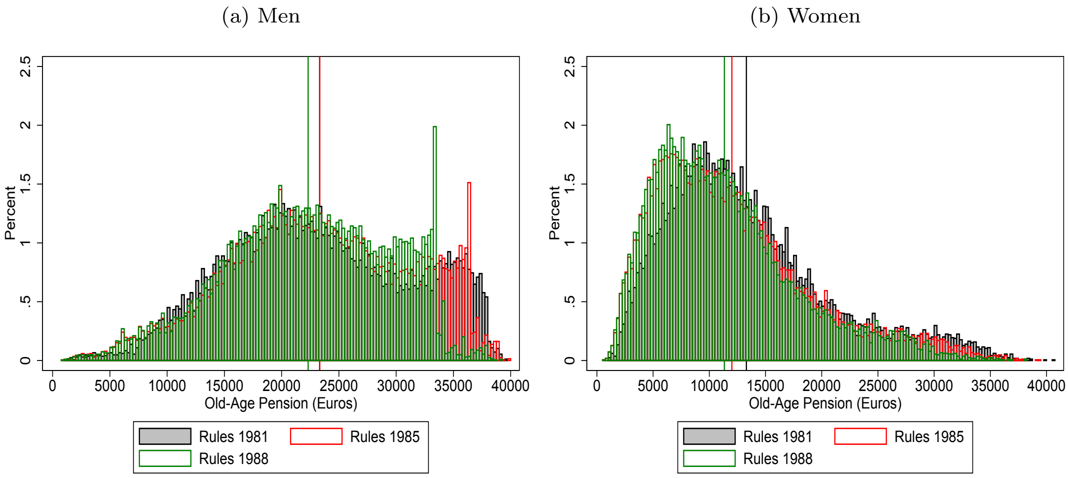
Sample Selection. My main sample consists of the male birth cohorts 1927-1928 and female birth cohorts 1932-1933. I exclude individuals who have worked in publicly-owned industries (public administration, public transportation, and education), as public sector workers are covered by a separate pension system with different eligibility rules. I further exclude self-employed individuals and individuals who have spent any time working in jobs defined as heavy labor, as they might be eligible for a special heavy labor pension. Lastly, I exclude individuals who are on DI before the reform, since they already left the labor force and are not affected by the reform.

4.2.1 Descriptive Evidence

Figure 11 plots the distribution of pension levels under the rules in 1981, 1985 and 1988. The figure shows that the increase of the assessment period from the last 5 to the last 10 to the last 15 years reduced pension levels on average but there is significant heterogeneity. Notably, men at the top of the pension distribution face a reduction in pension levels, while men at the lower end of the distribution are less affected. The effects for women are smaller because most women are at the lower end of the pension distribution. In my main analysis I focus on the change of

[22] DI benefits are calculated in the same way as old-age pensions, forward-looking individuals therefore have an incentive to claim DI benefits at earlier years before their assessment period increases by another year. This anticipation effect would be interesting to estimate. However, in 1984 relaxed DI eligibility criteria were introduced. This led to strong take-up of DI after age 55 and roll-out of these more lenient rules led to very strong trends in DI take-up across birth cohorts. This makes it hard to disentangle time trends from birth cohort effects. I therefore focus on the cohorts with the fixed rules and exploit the variation in an RDD, where differential trends across birth cohorts are less of a concern.

Figure 11: Pension Distribution pre/post 1985 and 1988 Reform



Notes: This figure plots the distribution of pension benefits under the rules in 1981, 1985 and 1988. To isolate the effect of the change in the pension calculation, I take the male birthcohort 1927 and the female birth cohort 1932 and simulate their potential pensions under the different regimes holding their retirement behavior fixed.

the assessment period from 10 to 11 years. Figure 12 plots the distribution of pension benefits with an assessment period of the last 10 versus the last 11 years. Men at the top of the pension distribution are strongly treated by the change in the assessment period. For women the effects are more uniform across the distribution and on average smaller. This makes it more difficult to precisely identify effects for women. The effects for men are stronger. I therefore focus on men in the main text, the results for women are in appendix D.

4.2.2 Empirical Strategy

I exploit the variation in pension levels induced by differential assessment periods across birth cohorts in an regression-discontinuity design (RDD). The running variable is date of birth. The assessment period for men born before January 1st 1928 is 10 years, for men born between 1/1/1928 and 12/31/1928 the assessment period is 11 years. I estimate the following regression

$$Y_i = \beta D_i + f_l(bdate_i)1\{bdate_i < 1/1/1928\} + f_r(bdate_i)1\{bdate_i \geq 1/1/1928\} + \varepsilon_i, \qquad (12)$$

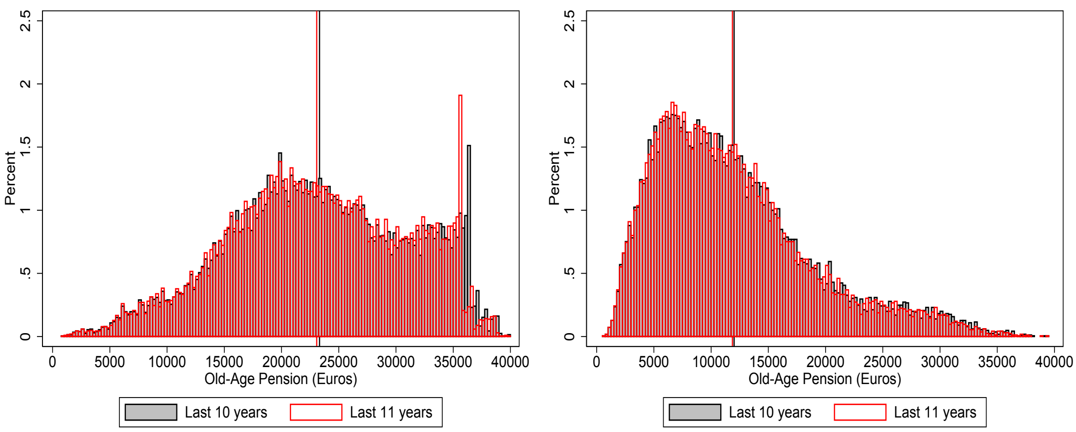
where $D_i = 1\{bdate_i \geq 1/1/1928\}$ is an indicator for individual i being born after January 1st 1928, $f_l(\cdot)$ and $f_r(\cdot)$ are flexible functions to capture trends in the outcome variable by date

Table 3: Phase in Pension Reform in 1988

Year	Men born	Women born	Assessment period	If more favorable
1988	until 1927	until 1932	120 months	132 months
	after 1928	after 1933	132 months	
1989	until 1927	until 1932	120 months	144 months
	in 1928	in 1933	132 months	144 months
	after 1929	after 1934	144 months	
1990	until 1927	until 1932	120 months	156 months
	in 1928	in 1933	132 months	156 months
	in 1929	in 1934	144 months	156 months
	after 1930	after 1935	156 months	
1991	until 1927	until 1932	120 months	168 months
	in 1928	in 1933	132 months	168 months
	in 1929	in 1934	144 months	168 months
	in 1930	in 1935	156 months	168 months
	after 1931	after 1936	168 months	
1992	until 1927	until 1932	120 months	180 months
	in 1928	in 1933	132 months	180 months
	in 1929	in 1934	144 months	180 months
	in 1930	in 1935	156 months	180 months
	in 1931	in 1936	168 months	180 months
	after 1932	after 1937	180 months	

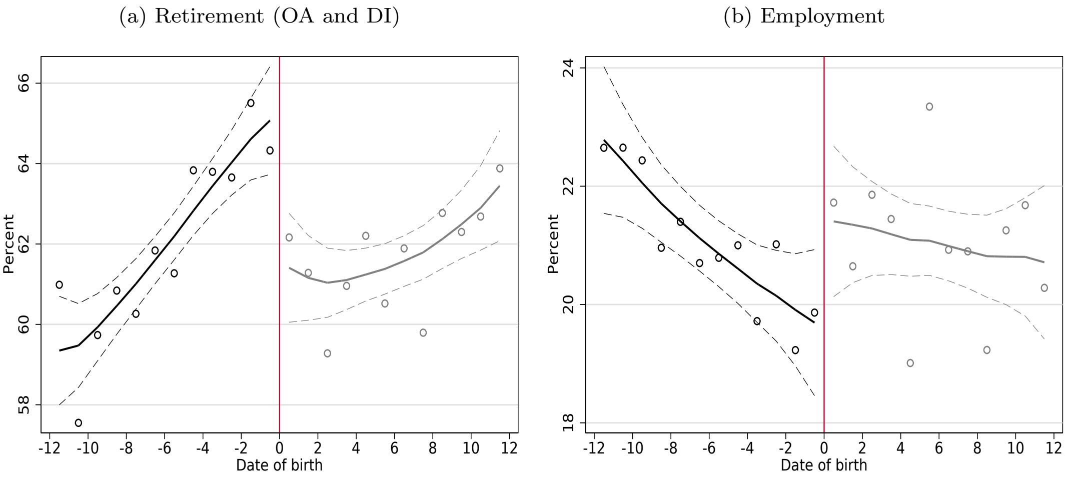
Figure 12: Pension Distribution with Assessment Period last 10 vs. last 11 years

(a) Men (b) Women



Notes: This figure plots the distribution of pension benefits with assessment period last 10 years versus last 11 years. To isolate the effect of the change in the assessment period, I take the male birthcohort 1927 and the female birth cohort 1932 and simulate their potential pensions under the different regimes holding their retirement behavior fixed. Hence, the shift in the distribution illustrates the mechanical effect of the reform. The vertical lines represent the means.

Figure 13: Retirement and Employment at Age 60



Notes: This figure shows the average retirement and employment rates at age 60. The fitted lines are local linear polynomials with a bandwidth of 8 months.

of birth. I am again interested in the total fiscal effect of the reform, the decomposition of this effect into the behavioral and mechanical part, and understanding the labor market responses of individuals.[23]

4.2.3 Results

Labor Market Responses. Figure 13 shows that the reduction in pension levels has significant effects on retirement take-up at age 60. Individuals to the left of the cutoff have an assessment basis of 10 years, individuals to the right of the cutoff have an assessment basis of 11 years. This leads to a reduction of around 1.25 percent in pensions on average. The lower pensions go hand-in-hand with a reduction of 2 percentage points in retirement take-up at age 60 (panel (a)) and to a less precisely estimated increase in employment of around 2 percentage points (panel (b)). There are no effects on UI or SI take-up. Even though treatment intensity is not very strong with a 1.25 percent reduction in pension levels at the cutoff, individuals react in their labor supply decision.

The increase in the assessment period from 10 to 11 years does not lead to a uniform cut in pension levels. The corresponding pension reduction varies by earnings history. Individuals with flat wage profiles are not treated by the change, while individuals with seniority wage profiles face

[23]Figure 36 in appendix D shows that the number of observations are not perfectly smooth around January for all birth cohorts. I discuss in appendix D why this seems not to be a problem for my RDD.

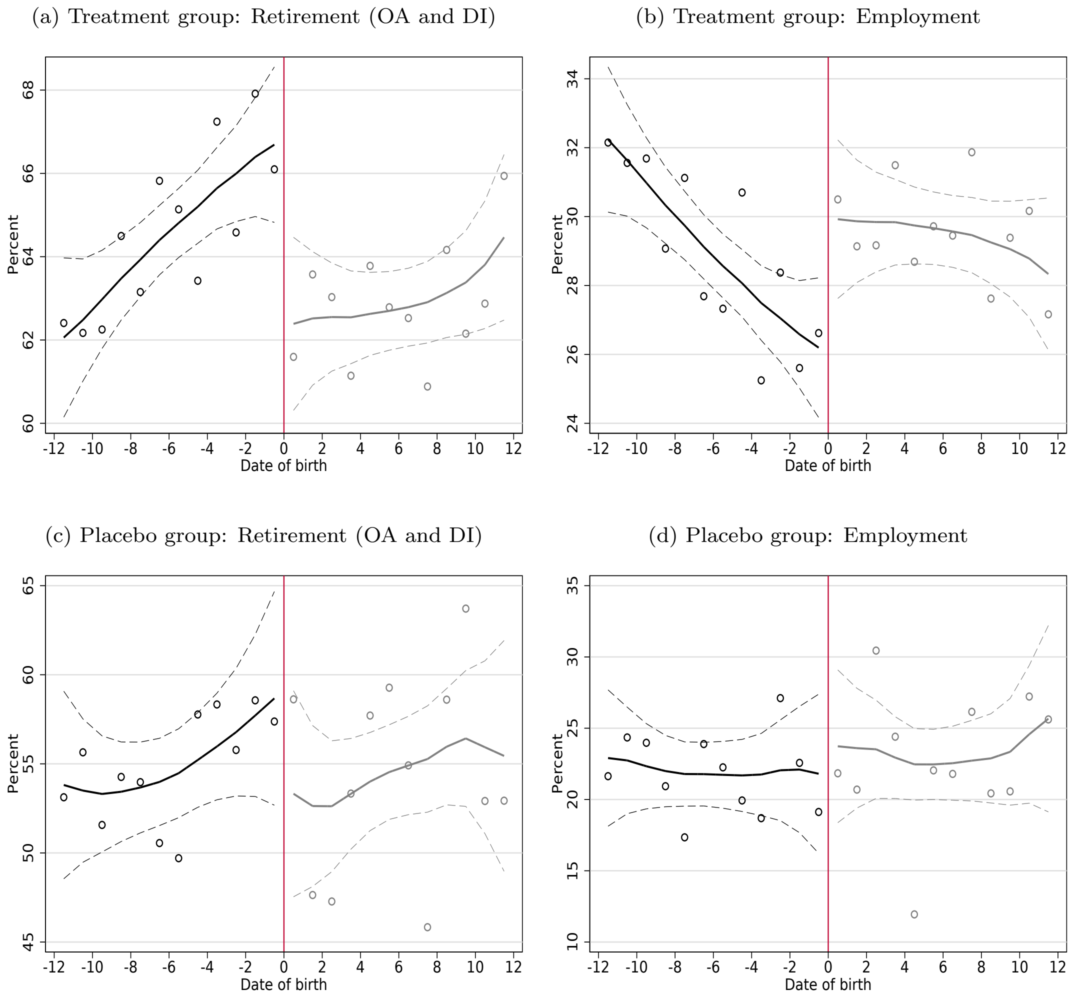
a reduction in pension levels due to the reform. This naturally creates placebo and treatment groups. To exploit this, I calculate for each individual the potential pension he would get with an assessment period of 10 and 11 years given his earnings history at age 59.5. Based on these potential pensions I define two groups: (i) a treatment group composed of individuals, who experience a reduction in pensions of at least 1.5% and (ii) a placebo group composed of individuals, who have roughly equal pensions in both regimes (change in pensions is between 0 and 0.25%). This definition is arbitrary and based on sample size considerations.[24]Figure 14 shows retirement and employment rates at age 60 for these two groups. The figure reveals that the retirement and employment effect we observe in the whole sample is driven by the treatment group. For the treatment group (individuals with a change in pensions of more than 1.5%), retirement at age 60 drops from 66% to 62% at the cutoff (panel (a)). This reduction in retirement rates by 4 percentage points is offset by a 4 percentage points increase in employment (panel (b)). Panels (c) and (d) show that there is no significant effect in the placebo group.

Figure 15 plots the RD estimates at each age between 50 and 85. There are no effects before age 59. Individuals just learn about the reform at age 59 and hence this is another placebo test for the validity of the RDD. There is a significant reduction in retirement at age 60 and a corresponding increase in employment. For ages 61 to 65 the point estimates still suggest a lasting effect on retirement and employment. However, these effects are not precisely estimated and not statistically different from zero. After the normal retirement age at 65, the point estimates are close to zero.

Fiscal Revenue Effect and Multiplier. Figure 15 already indicates that the policy change induced behavioral change. Figure 16 shows how these labor supply effects translate to fiscal revenue effects. Figure 16 plots the RD estimates for fiscal revenue at each age between 50 and 85 for men and women. The red area indicates the mechanical fiscal effect of the reform. Before the early retirement age there is no mechanical effect. There are also no effects on fiscal revenue as we would expect if the RD is valid (individuals did not yet know about the reform at these ages). Between the early and normal retirement age, the estimated total fiscal effect is larger than the mechanical fiscal effect. After the normal retirement age the total fiscal effect very closely follows the mechanical effect. The difference between the total fiscal effect and the mechanical fiscal effect is the behavioral fiscal effect. Hence, the figure corresponds to what one would expect from the

[24]Roughly 25% of all individuals experience a decline in pensions of more than 1.5%. Taking a higher cutoff value would significantly reduce the sample size.

Figure 14: Employment and Retirement at Age 60



(a) Treatment group: Retirement (OA and DI)

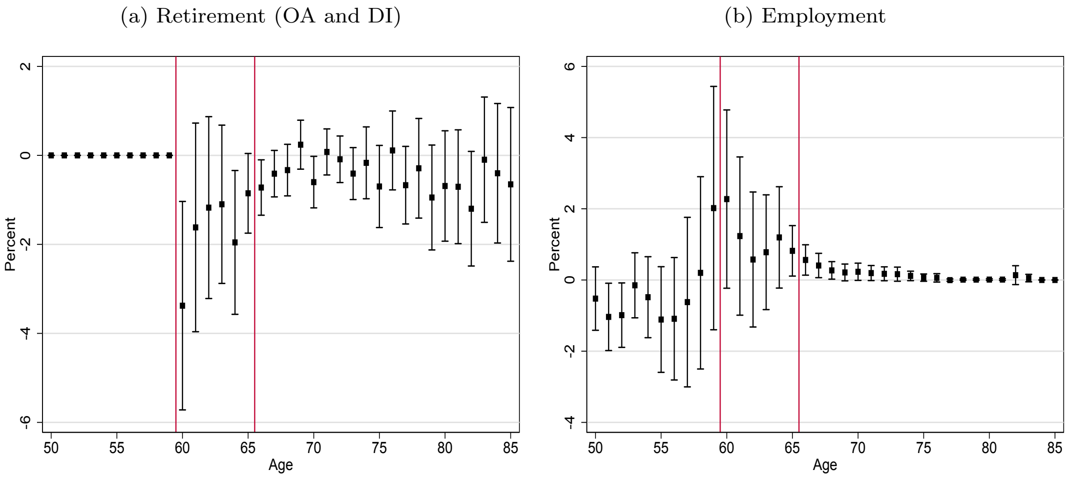
(b) Treatment group: Employment

(c) Placebo group: Retirement (OA and DI)

(d) Placebo group: Employment

Notes: This figure shows the average retirement and employment rates at age 60. The fitted lines are local linear polynomials with a bandwidth of 8 months. The treatement group is defined as individuals who potentially loose more than 1.5 percent in their OA pension if the assessment period is changed from 10 to 11 years. This potential loss is measured at age 59.5. The placebo group consists of individuals, who experience a loss of less than 0.25 percent.

Figure 15: RD Estimates by Age: Men

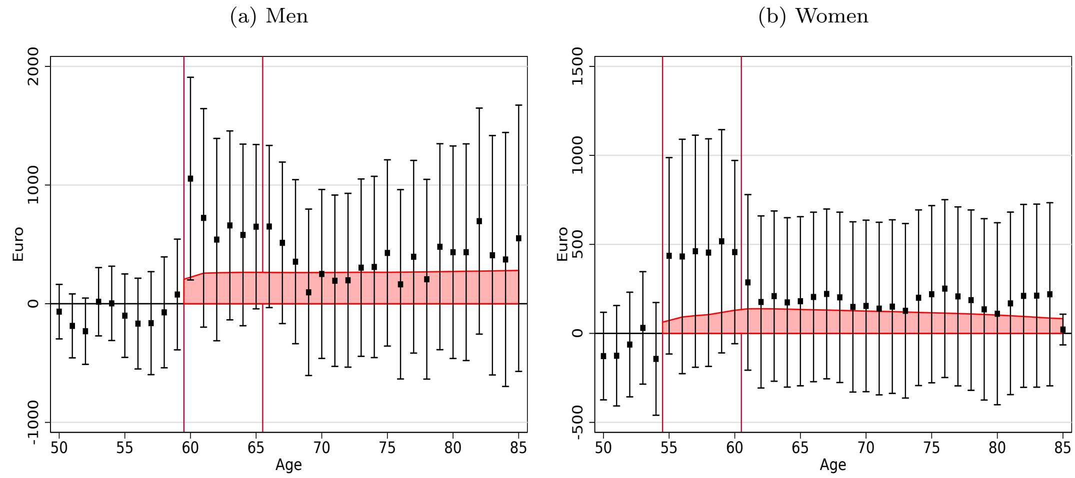


Notes: This figure plots the RD estimates at each age between 50 and 85 based on local linear regressions with a bandwith of 8 months.

labor supply responses. There is a behavioral fiscal effect from individuals delaying retirement and working longer. After the normal retirement age, when everyone is retired, only the mechanical effect of the reform is left. This is a consistent pattern for both men and women but the effects are not precisely estimated. However, this is not surprising since treatment intensity is relatively low with an average reduction in pension levels of 1.25 percent for men (with a median change of 1.5 percent) and 1.2 percent for women (with a median change of 1.2 percent).

Based on these estimates I construct the present value of the total fiscal effect and the present value of the mechanical fiscal effect at age 60. I discount the effects with an interest rate of two percent and then add up the estimates from age 60 to age 85. Above the normal retirement age I set the fiscal revenue estimates equal to the mechanical effect, since they are almost identical. Not doing this would lead to higher multipliers. Table 4 shows the present value of the total fiscal revenue effect, of the total mechanical fiscal effect and the corresponding multiplier. For men the multiplier of the reform is 1.48. For women the multiplier is 1.84. Table 7 in Appendix shows multipliers for different discount rates and for different bandwidths of the local linear regressions. For lower bandwidths and higher discount rates multipliers tend to be higher.

Figure 16: RD Estimates by Age: Fiscal Effect



Notes: This figure plots the RD estimates for fiscal revenue by age (black squares with 95 CI, based on local linear regressions with bandwidth 8 months). The red area indicates the mechanical fiscal effect of the reform. Between the ERA and NRA the estimated total fiscal effect is larger than the mechanical fiscal effect. After the NRA the total fiscal effect very closely follows the mechanical effect. Hence, there is a behavioral fiscal effect (difference btw. total and mechanical) between ERA and NRA. After the NRA, when everyone is retired, only the mechanical effect of the reform is left.

Table 4: Multiplier Benefit Generosity Reform 1988

Group	Fiscal Revenue Effect	Mechanical	Behavioral	Multiplier (1+B/M)
Men	7995	5408	2587	1.48
Women	4735	2578	2157	1.84

4.2.4 Welfare Implications

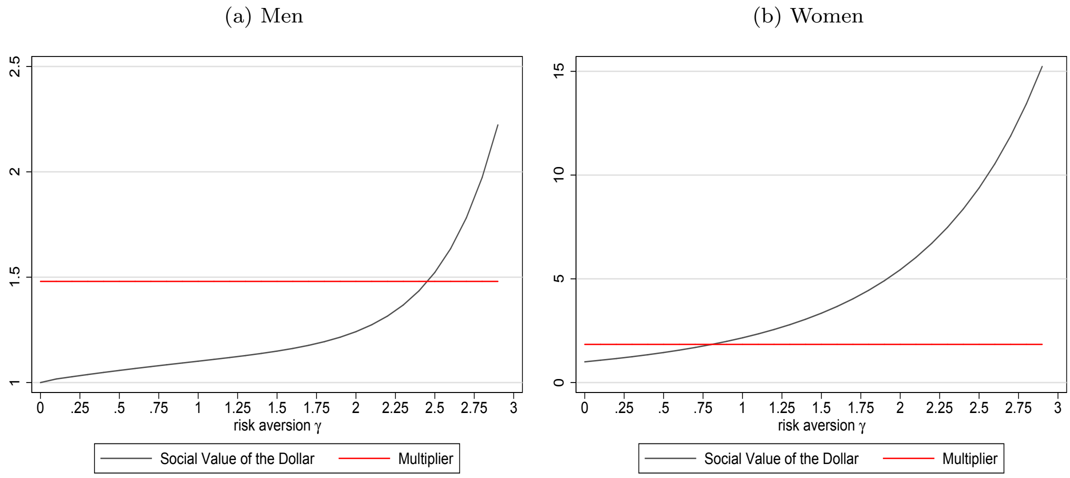
The relative comparison between the ERA reform and the pension level reform reveals a clear pattern. The multipliers of a reduction in pension levels are 50 percent larger than the multipliers of increasing the ERA. Additionally, the reduction in pension levels mostly affects individuals at the top of the income distribution while the ERA reform affects individuals across the income distribution. Based on the discussion in section 2, this implies for a planner with preferences for redistribution that the reduction in pension levels has a lower social value of the dollar compared to increasing the ERA because the pension level reform targets higher income individuals. Hence, a planner with preferences for redistribution clearly favors the pension level reform over the ERA reform, because it has higher multipliers and comes at a lower social cost (lower social value of the dollar).

The absolute welfare effect of the pension level reform is positive if the social value of the dollar is less than 1.48 for men and at less than 1.84 for women. To illustrate the magnitude of the social value of the dollar, I parameterize the model. In summary, I assume a standard isoelastic (CRRA) utility function and that the fiscal savings from the pension reform are lump-sum redistributed in the working age population, i.e. λ corresponds to the average marginal utility of consumption of working age individuals. Furthermore, I assume that consumption equals income since consumption cannot be observed in the data. This potentially overestimates the social value of the dollar. Working age individuals tend to save (consume less than their income), while retirees tend to dissave (consume more than their pension payments). Hence, assuming consumption equals income overestimates the marginal utility of consumption of retirees (numerator of social value of the dollar) and underestimates the marginal utility of consumption of working age individuals (denominator of social value of the dollar). Appendix A.5 provides a detailed discussion of this parametrization.

With these assumptions, I can directly calculate the social value of the dollar as a function of the risk aversion parameter γ. Figure 17 plots the social value of the dollar and the multiplier of the 1988 pension reform for men and women. Panel (a) shows that reducing pension levels of men is welfare-enhancing if risk aversion is below 2.4, i.e. the multiplier is larger than the social value of the dollar. Hence, this parametrization implies that the reform is welfare improving for reasonable values of risk aversion .[25] Panel (b) shows reducing pension levels of women is welfare-reducing if risk-aversion is above 0.75 even though women have a larger multiplier than men. The different welfare implications for men and women arise because the pension levels of women are significantly lower than the pension levels of men, which implies a much larger social value of the dollar for women. However, individual pension income of women in 1988 in Austria is probably a poor approximation for their consumption. Income at the household level would be a better predictor but cannot be observed in my data. For men individual income is a better approximation for household resources in that time.

[25] Estimates from the literature suggest that the coefficient of relative risk aversion is below 2, Chetty (2006b) finds an upper bound of $\gamma < 1.78$.

Figure 17: Welfare Effects of Reducing Pension Levels



Notes: This figure plots the social value of the dollar and the multiplier of the 1988 pension reform. The social value of the dollar is based on the parameterization of the utility function (additive separable CRRA utility function) as detailed in Appendix A.5. Reducing pension levels of men is welfare-enhancing (multiplier is larger than social value of the dollar) if risk aversion is below 2.4. For women, reducing pension levels is welfare-reducing if risk-aversion is above 0.75. The large difference between men and women arises because the pension levels of women are significantly lower than the pension levels of men, which implies a much larger social value of the dollar for women.

5 Conclusion

This paper shows that the behavioral fiscal multiplier is central for welfare evaluation of pension reforms. The behavioral fiscal multiplier, the total fiscal effect relative to the mechanical fiscal effect, can be readily estimated with reduced-form methods and provides the benchmark against which to judge the social value of the dollar. Exploiting a series of pension reforms in Austria, I find that increasing the early retirement age has a multiplier of 1 and reducing pension levels has a multiplier of 1.5. This implies that increasing the Austrian early retirement age has not been welfare-enhancing – unless one thinks that $1 in the hands of an early retiree has a lower social value than $1 in public funds. By contrast, reducing pension levels is welfare improving – provided that taking $1 away from a retiree is associated with a social loss smaller than $1.5. For a standard parametrization of the utility function I find that the social loss is smaller than 1.5 for reasonable values of risk aversion and hence reducing pension levels was welfare-improving. The low multiplier of increasing the early retirement age arises because the additional social security contributions paid by workers spending more time in employment are neutralized by additional expenditures on unemployment benefits. In contrast, reducing pension levels induced some workers to stay longer in employment without triggering substitution to other welfare benefits. My empirical analysis illustrated my framework in the Austrian context. Needless to say that the estimated multiplier depends on the particular context. If my framework is applied to reforms in other countries, the welfare implications can be very different depending on the labor market responses of affected workers.

The more general message of my analysis is twofold. First, the size of the behavioral multiplier crucially depends on the extent to which a particular pension reform generates spillovers to other pension reforms. Second, pension reforms generate high multipliers if older workers face a labor market with favorable job opportunities. For instance, increasing retirement ages does not generate large multipliers if older workers cannot find or keep their jobs. To some degree, labor market opportunities for older workers are a policy parameter. Hence, policy makers can influence multipliers and the effectiveness of reforms. Potential labor market policies to achieve this are active labor market policies targeting older unemployed individuals, providing financial incentives for firms to keep older workers employed (e.g. reduced social security contributions) or increased job protection to keep older workers in employment. The effectiveness of such policies is an important topic for future research, especially the role of firms seems understudied in this respect. Another important topic for future research is to relate the social value of the dollar to observable moments in the data. In the unemployment literature, there are exciting new approaches on this frontier (e.g. Landais and Spinnewijn (2019) and Hendren (2017)), which could also be applied in the retirement context.

References

Behaghel, Luc and David M. Blau, "Framing Social Security Reform: Behavioral Responses to Changes in the Full Retirement Age," *American Economic Journal: Economic Policy*, May 2012, *4* (4), 41–67. 1, A.6.3

Benveniste, L. M. and J. A. Scheinkman, "On the Differentiability of the Value Function in Dynamic Models of Economics," *Econometrica*, 1979, *47* (3), 727–732. A.1

Berkovec, James and Steven Stern, "Job Exit Behavior of Older Men," *Econometrica*, 1991, *59* (1), 189–210. 1

Bernheim, B. Douglas, "BEHAVIORAL WELFARE ECONOMICS," *Journal of the European Economic Association*, 2009, *7*, 267–319. 30

_ , "The Good, the Bad, and the Ugly: A Unified Approach to Behavioral Welfare Economics," *Journal of Benefit Cost Analysis*, 2016, *7* (1), 12–68. 30

Bernheim, B Douglas and Dmitry Taubinsky, "Behavioral public economics," Technical Report, National Bureau of Economic Research 2018. A.6.3, A.6.3

Brown, Jeffrey R., Arie Kapteyn, and Olivia S. Mitchell, "FRAMING AND CLAIMING: HOW INFORMATION-FRAMING AFFECTS EXPECTED SOCIAL SECURITY CLAIMING BEHAVIOR," *Journal of Risk and Insurance*, 2016, *83* (1), 139–162. A.6.3

Chetty, Raj, "A General Formula for the Optimal Level of Social Insurance," *Journal of Public Economics*, November 2006, *90* (10-11), 1879–1901. 1

_ , "A New Method of Estimating Risk Aversion," *American Economic Review*, December 2006, *96* (5), 1821–1834. 25

_ , "Moral Hazard versus Liquidity and Optimal Unemployment Insurance," *Journal of Political Economy*, April 2008, *116* (2), 173–234. 1

_ **and Amy Finkelstein**, "Social Insurance: Connecting Theory to Data," in Alan J. Auerbach, Raj Chetty, Martin Feldstein, and Emmanuel Saez, eds., *Handbook of Public Economics: Volume 5*, Elsevier, 2013, chapter 3, pp. 111–193. 1

Clausen, Andrew and Carlo Strub, "A General and Intuitive Envelope Theorem," Technical Report 2016. A.1, A.1

Cribb, Jonathan, Carl Emmerson, and Gemma Tetlow, "Signals matter? Large retirement responses to limited financial incentives," *Labour Economics*, 2016, *42*, 203 – 212. 1, A.6.3

Duggan, Mark, Perry Singleton, and Jae Song, "Aching to retire? The rise in the full retirement age and its impact on the social security disability rolls," *Journal of Public Economics*, 2007, *91* (7), 1327 – 1350. 1

Farhi, Emmanuel and Xavier Gabaix, "Optimal Taxation with Behavioral Agents," Working Paper 21524, National Bureau of Economic Research September 2015. A.6.3, A.6.3

Fink, Marcel, "Annual National Report 2009 Pension, Health and Long-term Care," Technical Report, asisp Network 2009. 11

French, Eric, "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour," *The Review of Economic Studies*, 04 2005, *72* (2), 395–427. 1

Goda, Gopi Shah, Matthew R Levy, Colleen Flaherty Manchester, Aaron Sojourner, and Joshua Tasoff, "The Role of Time Preferences and Exponential-Growth Bias in Retirement Savings," Working Paper 21482, National Bureau of Economic Research August 2015. A.6.3

Gustman, Alan L. and Thomas L. Steinmeier, "A Structural Retirement Model," *Econometrica*, 1986, *54* (3), 555–584. 1

_ **and** _ , "Effects of social security policies on benefit claiming, retirement and saving," *Journal of Public Economics*, 2015, *129*, 51 – 62. 1

Haller, Andreas, Stefan Staubli, and Josef Zweimüller, "Tightening Disability Screening Or Reducing Disability Benefits? Evidence and Welfare Implications," Technical Report, Working paper 2019. 17

Hendren, Nathan and Ben Sprung-Keyser, "A Unified Welfare Analysis of Government Policies," Working Paper 26144, National Bureau of Economic Research August 2019. 1

Hendren, Nathaniel, "Knowledge of Future Job Loss and Implications for Unemployment Insurance," *American Economic Review*, 2017, *107* (7), 1778–1823. Replication code and files. 3, 5

HV, "Jahresbericht der Österreichischen Sozialversicherungen," Technical Report, Hauptverband der Sozialverischerung 2018. 10

Imrohoroglu, Selahattin and Sagiri Kitao, "Social Security Reforms: Benefit Claiming, Labor Force Participation, and Long-Run Sustainability," *American Economic Journal: Macroeconomics*, July 2012, *4* (3), 96–127. 1

Iskhakov, Fedor, "Structural dynamic model of retirement with latent health indicator," *The Econometrics Journal*, 2010, *13* (3), S126–S161. 1

Kleven, Henrik, "Sufficient Statistics Revisited," Technical Report 2018. 1

Kolsrud, Jonas, Camille Landais, Peter Nilsson, and Johannes Spinnewijn, "The Optimal Timing of Unemployment Benefits: Theory and Evidence from Sweden," *American Economic Review*, April 2018, *108* (4-5), 985–1033. 1

Laitner, John and Dan Silverman, "Consumption, retirement and social security: Evaluating the efficiency of reform that encourages longer careers," *Journal of Public Economics*, 2012, *96* (7), 615 – 634. 1

Landais, Camille and Johannes Spinnewijn, "The Value of Unemployment Insurance," Working Paper, LSE 2019. 3, 5

Lee, David S, Pauline Leung, Christopher J O'Leary, Zhuan Pei, and Simon Quach, "Are Sufficient Statistics Necessary? Nonparametric Measurement of Deadweight Loss from Unemployment Insurance," Working Paper 25574, National Bureau of Economic Research February 2019. 1

Luenberger, D.G., *Optimization by Vector Space Methods* Professional Series, Wiley, 1997. 8

Manoli, Day and Andrea Weber, "Nonparametric evidence on the effects of financial incentives on retirement decisions," *American Economic Journal: Economic Policy*, 2016, *8* (4), 160–82. 1

Manoli, Dayanand S and Andrea Weber, "The Effects of the Early Retirement Age on Retirement Decisions," Working Paper 22561, National Bureau of Economic Research August 2016. 1

Mastrobuoni, Giovanni, "Labor supply effects of the recent social security benefit cuts: Empirical estimates using cohort discontinuities," *Journal of Public Economics*, 2009, *93* (11), 1224 – 1233. 1

Merkle, Christoph, Philipp Schreiber, and Martin Weber, "Framing and retirement age: The gap between willingness-to-accept and willingness-to-pay," *Economic Policy*, 09 2017, *32* (92), 757–809. A.6.3

Milgrom, Paul and Ilya Segal, "Envelope theorems for arbitrary choice sets," *Econometrica*, 2002, *70* (2), 583–601. A.1

Rust, John and Christopher Phelan, "How Social Security and Medicare Affect Retirement Behavior In a World of Incomplete Markets Author," *Econometrica*, 1997, *65* (4), 781–831. 1

Schmieder, Johannes F., Till von Wachter, and Stefan Bender, "The Effects of Extended Unemployment Insurance Over the Business Cycle: Evidence from Regression Discontinuity Estimates Over 20 Years," *Quarterly Journal of Economics*, May 2012, *127* (2), 701–752. 1

Seibold, Arthur, "Reference Points for Retirement Behavior: Evidence from German Pension Discontinuities," Technical Report 2019. 1, A.6.3

Shimer, Robert and Iván Werning, "Reservation Wages and Unemployment Insurance," *Quarterly Journal of Economics*, 2007, *122* (3), 1145–1185. 1

Staubli, Stefan, "The Impact of Stricter Criteria for Disability Insurance on Labor Force Participation," *Journal of Public Economics*, October 2011, *95* (9-10), 1223–1235. 17

_ **and Josef Zweimüller**, "Does Raising the Early Retirement Age Increase Employment of Older Workers?," *Journal of Public Economics*, 2013, *108*, 17–32. 1, 16

Stock, James H. and David A. Wise, "Pensions, the Option Value of Work, and Retirement," *Econometrica*, 1990, *58* (5), 1151–1180. 1

van der Klaauw, Wilbert and Kenneth I. Wolpin, "Social security and the retirement and savings behavior of low-income households," *Journal of Econometrics*, 2008, *145* (1), 21 – 42. The use of econometrics in informing public policy makers. 1

Zweimüller, Josef, Rudolf Winter-Ebmer, Rafael Lalive, Andreas Kuhn, Jean-Philipe Wuellrich, Oliver Ruf, and Simon Büchi, "Austrian Social Security Database," *Working Paper University of Zurich*, 2009, *401*, 1–80. 3.4

A Technical Appendix

A.1 Derivation Optimal Benefit Function

This section derives the formula for the optimal benefit function in (9). The derivation establishes the differentiability of the government's objective function and that the first order condition (6) holds for any optimal benefit function. The derivation exploits the powerful "differentiable sandwich lemma" from Clausen and Strub (2016), who generalize the envelope theorems in Milgrom and Segal (2002). For the derivation, I impose three technical assumptions. First, I assume that the utility function $U_i(C, \Pi, X)$ is twice continuously differentiable in consumption $c(x^t)$ with $\frac{\partial U_i(.)}{\partial c_t(x_t)} > 0$ and $\frac{\partial U_i(.)}{\partial c_t(x_t)} \to \infty$ for $c_t(x_t) \to 0$ and $\frac{\partial^2 U_i(.)}{\partial c_t(x_t)^2} \leq 0$. Second, I assume the existence of a solution to the agent's problem that can be represented by a Lagrangian. Third, I assume the planner's budget constraint is differentiable in b.

We can solve agent i's optimization problem in (1)-(3) with the following Lagrangian

$$\begin{aligned} \mathcal{L}_i(C, \Pi, \gamma, \eta) \quad = \quad & U_i(C, \Pi, X) \\ & -E\left[\sum_{t=0}^{T-1} \gamma_{i,t+1}\left(a_{t+1} - (1+r_t)a_t - y_t(x^t) + \tau(x^t) - b(x^t) + c_t(x^t) + q(p_t(x^t))\right)\right] \\ & -E\left[\sum_{t=0}^{T-1} \eta_{i,t+1}\left(s_{t+1} - f\left[s^t, \pi_t(x^t), \varepsilon_t\right]\right)\right]. \end{aligned} \tag{13}$$

Let a solution to this problem be denoted by C_i^*, Π_i^* and the corresponding Lagrange multipliers be denoted by $\gamma_i^* = \{\gamma_{i,j+1}\}_{j=0}^{T-1}$ and $\eta^* = \{\eta_{i,j+1}\}_{j=0}^{T-1}$. The optimal consumption choice solves

$$\frac{\partial U_i(C, \Pi, X)}{\partial c_t(x_t)} = \gamma_{i,t+1}. \tag{14}$$

The government's optimization problem is to choose functions $b(\cdot)$.[26] We can express the planner's objective function as $W(b) = \int_{i \in I} \omega_i V_i(b) = \int_{i \in I} \omega_i \mathcal{L}_\text{i}(C(b), \Pi(b), \gamma(b), \eta(b)) di$, where $C(b), \Pi(b)$ denote the optimal choices of the agent for a given benefit function b and $\gamma(b), \eta(b)$ are the corresponding Lagrange multipliers. The Lagrangian of the planner's problem in (4) and (5) is given by

$$\mathcal{L}_{\text{SP}}(b) = \int_{i \in I} \omega_i \mathcal{L}_\text{i}(C(b), \Pi(b), \gamma(b), \eta(b)) di + \lambda\left(G(b) - \bar{G}\right). \tag{15}$$

[26]For notational convenience I suppress the tax function $\tau(\cdot)$. Optimization of the tax function would add an additional first order condition identical with the one for the benefit function, just with opposite sign.

Suppose b^* is an optimal solution to the planner's problem. To establish differentiability of the planner's objective function I construct a lower and upper support function and then apply the differentiable sandwich lemma. As an upper support function I take $C(b) = W(b^*)$, where $W(b^*)$ is the welfare associated with the optimal benefit function b^*. By definition we have $W(b^*) \geq W(b)\ \forall b$.This upper support function is therefore simply a constant and $C'(b) = \delta W(b^*; h) = 0$. The lower support function is given by $L(b) = \int_{i\in I} \omega_i \mathcal{L}_{\mathrm{i}}(C(b^*), \Pi(b^*), \gamma(b^*), \eta(b^*))di + \lambda\left(G(b) - \bar{G}\right)$. $\mathcal{L}_{\mathrm{i}}(C(b^*), \Pi(b^*), \gamma(b^*), \eta(b^*))$ is the indirect utility of the agent, holding his behavior for benefit function b^* fixed. This corresponds to the idea of a lazy decision maker that uses a completely unresponsive policy rule as in Benveniste and Scheinkman (1979). We have $\mathcal{L}_{\mathrm{i}}(C(b^*), \Pi(b^*), \gamma(b^*), \eta(b^*)) \leq \mathcal{L}_{\mathrm{i}}(C(b), \Pi(b), \gamma(b), \eta(b))\ \forall b$. Therefore, $L(b)$ is a lower support function. Since I assume that the planner's budget constraint is differentiable, $L(b)$ is differentiable and the derivative is given by

$$L'(b) = \int_{i\in I} \omega_i E\left[\sum_{t=0}^{T-1} \gamma^*_{i,t+1} h(x_t)\right] di + \lambda \delta G(b^*; h) \tag{16}$$

By the differentiable sandwich lemma of Clausen and Strub (2016), we have $\delta W(b^*; h) = C'(b) = L'(b)$. Using $C'(b) = 0$ and (16) we get

$$\delta W(b^*; h) = \int_{i\in I} \omega_i E\left[\sum_{t=0}^{T-1} \gamma^*_{i,t+1} h(x_t)\right] di + \lambda \delta G(b^*; h) = 0. \tag{17}$$

Using (14) and

$$\delta G(b; h) = \sum_{t=0}^{T-1} \left[1 + r_t\right]^{-t} E\left[-h(x_t)\right] + \int \sum_{t=0}^{T-1} \int \left[1 + r\right]^{-t} \left[-b(x^t)\right] \delta\mu(x^t; h) dx^t di \tag{18}$$

yields the result in (9).

A.2 Lower Bound for the Social Value of the Dollar

To derive a lower bound for the social value of the dollar we first need to bound λ. For this, suppose that the tax system is such that we cannot improve welfare through lump-sum transfers, i.e. equation (9) holds for all lump-sum transfers $\overline{h}$.[27] In this case, we have

$$\lambda = \frac{\frac{1}{M(\overline{h})} \int \omega_i E_i \left[\sum_{t=0}^{T-1} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)} \overline{h} \right] di}{1 + \frac{B(\overline{h})}{M(\overline{h})}}. \tag{19}$$

Moreover, if we assume that for lump-sum transfers the behavioral response is small (income effects are small), i.e. $1 + \frac{B(\overline{h})}{M(\overline{h})} \approx 1$, we have

$$\lambda = \frac{1}{M(\overline{h})} \int \omega_i E_i \left[\sum_{t=0}^{T-1} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)} \overline{h} \right] di \tag{20}$$

and λ corresponds to the average marginal utility of consumption in the population. If there are income effects, individuals will adjust their labor supply in response to the transfers and we have $1 + \frac{B(\overline{h})}{M(\overline{h})} \geq 1$. As a consequence we have

$$\lambda \leq \frac{1}{M(\overline{h})} \int \omega_i E_i \left[\sum_{t=0}^{T-1} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)} \overline{h} \right] di. \tag{21}$$

For the social value of the dollar, this implies

$$\frac{\frac{1}{M(h)} \int \omega_i E_i \left[\sum_{t=0}^{T-1} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)} h(x^t) \right] di}{\lambda} \geq \frac{\frac{1}{M(h)} \int \omega_i E_i \left[\sum_{t=0}^{T-1} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)} h(x^t) \right] di}{\frac{1}{M(\overline{h})} \int \omega_i E_i \left[\sum_{t=0}^{T-1} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)} \overline{h} \right] di}, \tag{22}$$

where $\overline{h}$ is the lump-sum transfer, which would redistribute the mechanical effect from the reform $M(h)$ equally across all agents and periods and hence we have $M(h) = M(\overline{h})$. If we think that the marginal utility of individuals affected by the reform is larger than the average marginal utility in the population and assume equal welfare weights for all individuals, then

$$\text{Social Value of the Dollar} \geq \frac{\frac{1}{M(h)} \int E_i \left[\sum_{t=0}^{T-1} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)} h(x^t) \right] di}{\frac{1}{M(\overline{h})} \int E_i \left[\sum_{t=0}^{T-1} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)} \overline{h} \right] di} \geq 1. \tag{23}$$

For instance, this is the case if we assume that the utility function is concave in consumption (falling marginal utility) and that the reform targets individuals, who have below average consumption. These are all reasonable assumptions for pension reforms and hence a natural lower bound for the social value of the dollar is 1 dollar.

[27] Actually, it is sufficient to assume that increasing lump-sum taxes is not welfare-enhancing. In this case, we have $\lambda \leq \frac{\frac{1}{M(\overline{h})} \int \omega_i E_i \left[\sum_{t=0}^{T-1} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)} \overline{h} \right] di}{1 + \frac{B(\overline{h})}{M(\overline{h})}}$ for all lump-sum taxes $\overline{h}$ and the argument still goes through.

The assumption of equal welfare weights is not crucial. The argument goes through if the welfare weights are such that a transfer to retirees is socially more valuable than a transfer to all individuals.

A.3 Comparison to the Baily-Chetty Formula (Optimal Unemployment Insurance)

The famous Baily-Chetty formula for optimal unemployment insurance is a special case of my formula in equation (9). The stylized Baily-Chetty formula is given by

$$\frac{v'(c^u)}{u'(c^e)} = 1 + \varepsilon$$

where $\varepsilon = \frac{\partial D}{\partial b}\frac{b}{D}$, denotes the elasticity of unemployment duration D with respect to unemployment benefit generosity b.

The RHS of this formula exactly corresponds to my RHS term in (9). By definition we have $\varepsilon = \frac{\partial D}{\partial b}\frac{b}{D}$. Since individuals can only respond by adjusting their search effort and this affects fiscal expenditures through altered unemployment duration ∂D, the behavioral fiscal effect is given by $B = \partial D * b$. The mechanical fiscal effect is given by $M = \partial b * D$. Hence, the elasticity captures the behavioral over mechanical fiscal effect in this model, i.e. $\varepsilon = \frac{\partial D}{\partial b}\frac{b}{D} = \frac{B}{M}$.

In this stylized job search model individuals have a utility function for being employed $u(c^e)$ and a utility function for being unemployed $v(c^u)$. Both utility functions only depend on consumption levels and taxes are lump-sum. The cost of taxation λ is simply the marginal utility of the tax-payer and given by $u'(c^e)$. The direct welfare effect of changing unemployment benefit generosity in this model is given by $E_h\left[\frac{\partial U_i(C,\Pi,X)}{\partial c_t(x_t)}\right] = \frac{1}{\Delta M(h)}\int \omega_i E_i\left[\sum_{t=0}^{T-1}\frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)}h(x^t)\right]di = \frac{1}{D\partial b}Dv'(c^u)\partial b = v'(c^u)$. Hence, my formula applied to this stylized model exactly delivers the Baily-Chetty formula.

A.4 Multiple Generations

The model in the main text treats pensions systems as if each generation pays for their own pensions during their working life. However, an important aspect of pensions is inter-generational redistribution. I show that even with multiple generations it is key to estimate the multiplier of a pension reform in one generation. To introduce multiple generations, let there be a stream of generations indexed by $j \in \{1, 2, \ldots\}$ and each generation consists of a continuum of agents indexed by $i \in I_j$. Agents in each generation j solve the same maximization problem as in equations (1)-(3), but the constraints and objective functions can be generation specific.

Overlapping Generations: In the classical overlapping generations setup, the currently young generation $j+1$ finances the pensions of the currently old generation j. The problem of the planner is to choose benefit and tax functions for each generation j. The planner puts a welfare weight g_j on each generation j and solves the following problem

$$\max_{\{b_j(\cdot),\tau_j(\cdot)\}_{j\geq 1}} W(b,\tau) = \sum_{j\geq 1} g_j \int_{I_j} \omega_i V_i di \tag{24}$$

subject to

$$\sum_{t\geq 0} [1+r_t]^{-t} E_{j+1}\left[\tau_{j+1}(x^t)\right] - \sum_{t\geq 0} [1+r_t]^{-t} E_j\left[b_j(x^t)\right] \geq 0 \; \forall j \in \{1,2,\ldots\}. \tag{25}$$

For each generation there is a separate budget restriction requiring that the present value of benefit payments to the currently old generation equals the present value of the tax revenues from the young generation. The optimal benefit function of generation j then solves

$$\frac{g_j \frac{1}{M(h)} \int_{I_j} \omega_i E_i \left[\sum_{t\geq 0} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)} h(x^t)\right] di}{\lambda_{j+1}} = 1 + \frac{B_j(h)}{M_j(h)} \tag{26}$$

where $B_j(h)$ is the behavioral fiscal effect of generation j over their life cycle and $M_j(h)$ is the mechanical fiscal effect of generation j. Hence, the RHS of this optimality condition (26) is identical to the RHS of optimality condition (9) in the single generation setup. The only difference to the single generation version is the interpretation of the LHS. The LHS measures the social value of a dollar in the hand of the currently old generation j relative to the social value of a dollar in the hand of the currently young generation $j+1$. To make this more transparent, assume that the planner redistributes the savings from the pension reform with lump-sum transfers to generation $j+1$. In this case, the optimal benefit function solves

$$\frac{g_j}{g_{j+1}} \frac{\frac{1}{M(h)} \int_{I_j} \omega_i E_i \left[\sum_{t\geq 0} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)} h(x^t)\right] di}{\int_{I_{j+1}} \omega_i E_i \left[\sum_{t\geq 0} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)}\right] di} = 1 + \frac{B_j(h)}{M_j(h)}. \tag{27}$$

General Setup: The budget restriction in the overlapping generations framework above is too restrictive in the sense that the planner can smooth expenditures and revenue across multiple generations. In this more flexible formulation the planner solves

$$\max_{b_j(\cdot),\tau_j(\cdot)} W(b,\tau) = \sum_{j\geq 1} g_j \int \omega_i V_i di \tag{28}$$

$$G(b,\tau) = \sum_{j\geq 1}\sum_{t\geq 0} [1+r_t]^{-t} E\left[\tau_j(x^t) - b_j(x^t)\right] \geq 0. \tag{29}$$

The optimal benefit formula then solves

$$\frac{g_j \frac{1}{M(h)} \int_{I_j} \omega_i E_i \left[\sum_{t\geq 0} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)} h(x^t)\right] di}{\lambda} = 1 + \frac{B_j(h)}{M_j(h)} \tag{30}$$

and is therefore identical to the optimal benefit formula in the main text (except that there is the additional welfare weight g_i for the generation on the LHS). Importantly, in the multiple generation setup we also need to estimate the multiplier of the reform within one generation in the same way as in the one generation model in the main text.

A.5 Parametrization Model

The social value of the dollar consists of three parts: (i) the marginal utility of consumption, (ii) welfare weights and (iii) the social value of public funds λ. Hence, I need to make assumptions on these three margins.

First, I parameterize the utility function. I assume that utility is additively separable over time and additively separable in consumption and the other choices. I denote the utility of consumption in period t by $u(c_t)$ and the utility of the other choices by $\psi(\Pi, X)$. Hence, expected life-time utility is given by

$$U_i(C,\Pi,X) = E\left[\sum_{t=0}^{T-1} \left(\beta^t u(c_t)\right) - \psi(\Pi,X)\right]. \tag{31}$$

Moreover, I assume that the utility function is isoelastic, i.e.

$$u(c_t) = \begin{cases} \frac{c_t^{1-\gamma}-1}{1-\gamma} & \gamma \geq 1 \,\&\, \gamma > 0 \\ ln(c_t) & \gamma = 1 \end{cases}, \tag{32}$$

where γ denotes the coefficient of relative risk aversion.

Second, I assume a utilitarian planner, i.e. every individual gets the same welfare weight $\omega_i = 1\,\forall i$.

Third, I need to pin down the social value of public funds. Following the discussion in Appendix A.2, I assume that welfare cannot be improved through lump-sum transfers to working age individuals (younger than the ERA). This implies that

$$\lambda = \frac{1}{1+\frac{B(\bar{h})}{M(\bar{h})}} \int E_i \left[\sum_{t=0}^{ERA} \left(\beta^t u'(c_t)\bar{h} \right) \right] di \tag{33}$$

holds for any small lump-sum transfer $\bar{h}$. If we further assume that income effects are small, i.e. $1+\frac{B(\bar{h})}{M(\bar{h})} \approx 1$, we get that

$$\lambda \approx \int E_i \left[\sum_{t=0}^{ERA} \left(\beta^t u'(c_t) \right) \right] di. \tag{34}$$

The social value of the dollar is then given by

$$\text{Social Value of the Dollar} = \frac{\frac{1}{M}\int E_i \left[\sum_{t=ERA}^{T-1} \left(\beta^t u'(c_t) h_{it} \right) \right] di}{\int E_i \left[\sum_{t=0}^{ERA} \left(\beta^t u'(c_t) \right) \right] di} \tag{35}$$

where h_{it} is the reduction in individual i's pension in period t because of the reform and M is the mechanical effect of the reform (the sum of h_{it}).

In my data, consumption is not observable. As an approximation I use income instead of consumption. Hence, for retirees I set consumption equal to pension payments, i.e. $c_t = b_t$. For working age individuals I set consumption equal to labor income, i.e. $c_t = w_t$. This potentially overestimates the social value of the dollar. Working age individuals tend to save, i.e. $c_t \leq w_t$, and hence $u'(w_t) \leq u'(c_t)$. Retirees tend to dissave, i.e. $c_t \geq b_t$, and hence $u'(b_t) \geq u'(c_t)$. Therefore, we expect that

$$\frac{\frac{1}{M}\int E_i \left[\sum_{t=ERA}^{T-1} \left(\beta^t u'(c_t) h_{it} \right) \right] di}{\int E_i \left[\sum_{t=0}^{ERA} \left(\beta^t u'(c_t) \right) \right] di} \leq \frac{\frac{1}{M}\int E_i \left[\sum_{t=ERA}^{T-1} \left(\beta^t u'(b_t) h_{it} \right) \right] di}{\int E_i \left[\sum_{t=0}^{ERA} \left(\beta^t u'(w_t) \right) \right] di} \equiv \text{USVD} \tag{36}$$

holds and that my implementation provides an upper bound for the social value of the dollar (USVD). Lastly, I assume no discounting $\beta = 1$.This provides a further upper bound on the social value of the dollar (since the numerator of the social value of the dollar would be relatively smaller with discounting).

Implementation for 1988 Pension Level Reform. To implement the upper bound on the social value of the dollar (USVD as defined in (36)) for the 1988 pension reform, I calculate the mechanical loss h_{it} for each individual in the control group (men born in 1927 and women born in 1932) by calculating their old-age pensions with an assessment period of 10 and 11 years and then take the difference between the two.

Based on the CRRA specification I calculate the marginal utility of each individual for different values of γ, multiply it by h_{it} and take the average across individuals and periods. This average is then divided by the mechanical effect M (the average of h_{it} over individuals and periods). For the denominator, I take the earnings distribution of working-age individuals in the year 1988 and calculate the average marginal utility of consumption for this earnings distribution for different values of γ.

A.6 Extensions

The insight that the multiplier of a reform is central for welfare evaluation relies on three critical assumptions. First, I assume that individuals optimize a consistent utility function and the social planner respects the preferences of the individuals. Second, I assume that reforms are small. Third, I abstract from general equilibrium effects and externalities of pensions reforms. I discuss here how deviations from these three assumptions change the welfare evaluation. In general, any deviation from these three assumptions leads to an additional term on the LHS of equation (9) but does not change the RHS. Hence, in any case we need to know the multiplier of the reform.

A.6.1 General Equilibrium Effects

Allowing for price changes does not alter the logic of the welfare evaluation. With perfect competition the price adjustments only have a direct welfare effect through the mechanical effect of the price changes. With general equilibrium effects, the optimal benefit function formula becomes

$$\frac{\frac{1}{M(h)}\int_{I_j}\omega_i E_i\left[\sum_{t\geq 0}\frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)}h(x^t)\right]di}{\lambda}+\gamma=1+\frac{B(h)}{M(h)} \tag{37}$$

where $\gamma=\frac{\frac{1}{M(h)}\int_{I_j}\omega_i E_i\left[\sum_{t\geq 0}\frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)}dP(x^t)\right]di}{\lambda}$ and $dP(x^t)$ measures the price change due to the reform. Furthermore, $B(h)$ is evaluated at the new prices, while $M(h)$ is evaluated at the old prices. The sign of γ depends on the price changes. The general equilibrium effects of pension reforms we might worry about are changes in wages and interest rates. If a pension reform induces individuals to retire later and save more we might expect that interest rates fall (since there is more capital). This affects capital returns of individuals and the planner's budget constraint through the discounting (this is captured in the behavioral fiscal effect $B(h)$). The effect on wages is not a priori clear. It could be that the additional capital makes workers more productive and therefore wages increase. It could also be that the additional workers in the market reduce wages (increased labor supply). In the context of Austria it is unlikely that a pension reform affects interest rates, since Austria is a small open economy and does not even set their own monetary policy.

A.6.2 Non-marginal Changes

The optimal benefit formula only holds for marginal changes. However, real-world reforms are never truly marginal. In this section I discuss the welfare effect of discrete changes. Suppose there are two different benefit functions $b_0(\cdot)$ and $b_1(\cdot)$, where $b_1(x^t) = b_0(x^t) + h(x^t)$. The welfare effect of the discrete change from $b_0(\cdot)$ to $b_1(\cdot)$ is given by

$$\begin{aligned} \Delta W &= W(b_1, \tau) - W(b_0, \tau) && (38) \\ &= \int \omega_i \left[V_i(b_1) - V_i(b_0)\right] di + \lambda \left(G(b_1, \tau) - G(b_0, \tau)\right) && (39) \end{aligned}$$

where I assumed that $\lambda_0 = \lambda_1 = \lambda$. $G(b_1, \tau) - G(b_0, \tau)$ measures the total fiscal effect of moving from benefit function $b_0(\cdot)$ to $b_1(\cdot)$. Define $G(b_k) = \int \sum_{t\geq 0} [1 + r_t]^{-t} \int [(\tau(x) - b_k(x)) \mu_k(x)] \, dxdi$, where $b_k(x)$ denotes the benefit payment for state history x in regime $k = \{0, 1\}$ and $\mu_k(x)$ is the distribution of state histories in regime k. The state distribution $\mu_k(x)$ depends on the behavior of individuals. The total fiscal effect can then again be decomposed into a behavioral and mechanical effect:

$$G(b_1) - G(b_0) = \int \sum_{t\geq 0} [1 + r_t]^{-t} \int \left[\underbrace{\Delta b_0(x)\mu_1(x)}_{\text{mechanical}} + \underbrace{(\tau(x) - b_0(x)) \, \Delta\mu_0(x)}_{\text{behavioral}} \right] dxdi \quad (40)$$

where $\Delta b_0(x) := b_1(x) - b_0(x)$ and $\Delta\mu_0(x) := \mu_1(x) - \mu_0(x)$. The mechanical effect is now only slightly differently defined as

$$\Delta M(h) = \int \sum_{t\geq 0} [1 + r_t]^{-t} \int \Delta b_0(x)\mu_1(x) dxdi. \quad (41)$$

The mechanical fiscal effect is now evaluated at the post reform state distribution, but this is still straightforward to implement empirically. The behavioral fiscal effect is given by

$$\Delta B(h) = \int \sum_{t\geq 0} \left[1 + r_t\right]^{-t} \int \left(\tau(x) - b_0(x)\right) \Delta\mu_0(x) dx di. \tag{42}$$

For discrete changes the envelope theorem fails and behavioral adjustments have first order welfare effects. The direct welfare effect is given by

$$\begin{aligned} \int \omega_i \left[V_i(b_1) - V_i(b_0)\right] di &= \int \omega_i \left[V_i(b_0 + h) - V_i(b_0)\right] di \\ &= \int \omega_i \left[\delta V_i(b_0; h) + \frac{1}{2!}\delta^2 V_i(b_0; h, h) + \ldots + \frac{1}{(k-1)!}\delta^{k-1} V_i(b_0; h, \ldots, h) + R_k\right] di \end{aligned} \tag{43}$$

where the second line uses a Taylor approximation and the remainder term is $R_k = \frac{1}{(k-1)!}\int_0^1 (1-t)^{k-1}\delta^k V_i(b_0 + th; h, \ldots, h)dt$. Using $\delta V_i(b_0; h) = E_i\left[\sum_{t\geq 0} \frac{\partial U_i C, \Pi, X)}{\partial c_t(x^t)} h(x^t)\right]$ we can write the discrete welfare change as

$$\Delta W = \frac{\frac{1}{\Delta M(h)} \int \omega_i E_i \left[\sum_{t=0}^{T-1} \frac{\partial U_i(C,\Pi,X)}{\partial c_t(x^t)} h(x^t)\right] di}{\lambda} + \gamma - \left(1 + \frac{\Delta B(h)}{\Delta M(h)}\right) \tag{44}$$

where $\gamma = \frac{1}{\lambda \Delta M(h)} \left(\frac{1}{2!}\delta^2 V_i(b_0; h, h) + \ldots + \frac{1}{(k-1)!}\delta^{k-1} V_i(b_0; h, \ldots, h) + R_k\right)$. With non-marginal changes the social value of the dollar term has an additional component γ, since behavioral adjustments have first order welfare effects. Nevertheless, the multiplier of the reform is still a key ingredient for welfare evaluation.

A.6.3 Behavioral Biases

A recent and growing literature documents behavioral elements in retirement decisions (Behaghel and Blau, 2012; Goda et al., 2015; Brown et al., 2016; Cribb et al., 2016; Merkle et al., 2017; Seibold, 2019). This evidence raises the question how the welfare evaluation changes in presence of behavioral biases. Generally, behavioral biases induce a potential conflict between the planner's and the agents objectives. The behavioral welfare literature makes the distinction between decision utility and true or "experienced" utility. If the two do not coincide, agents do not maximize their true utility. As a consequence the logic of the envelope theorem fails and behavioral responses have first order welfare effects. Hence, there is an additional term in the direct welfare effect that measures the bias correction of behavioral responses. This is analogous to the effects of behavioral elements in optimal tax frameworks as in Farhi and Gabaix (2015) and Bernheim and Taubinsky (2018). The crucial question in models with behavioral elements is: What is the true utility and consequently what is the right welfare criterion? This discussion is interesting and important. However, this paper makes no progress on that frontier. I simply assume that true preferences exist and are known to the planner. My analysis here follows Farhi and Gabaix (2015) and I simply adjust notation to my setup.

Let $U(\theta)$ denote the true or "experienced" life-time utility for choices θ.[28] However, choices are no longer necessarily made by maximizing the true utility function $U(\theta)$. Agents could maximize another (biased) decision utility function or fail to maximize completely. The choice function θ incorporates all these potential behavioral aspects. As in Farhi and Gabaix (2015), the only restriction I impose is that the agent's budget restriction is binding. That is, $K(b,I) := \sum_{t=0}^{T-1} [1+r_t]^{-t} E[y_t(x^t) - \tau(x^t) + b(x^t) - c_t(x^t) - q(p_t(x^t))] - I = 0$, where I denotes the initial endowment. The indirect utility function is denoted by $V(b,I) = U(\theta(b,I))$. The government's objective function is then given by

$$W(b) = V(b,I) + \lambda\left(G(b) - \bar{G}\right)$$

where $G(b)$ is government revenue and defined as in the standard model.

A change of the benefit function with increment h has a welfare effect of $\delta W(b;h) = \delta V(b,I;h) + \lambda \delta G(b;h)$ and the optimal benefit function sets $\delta W(b;h) = 0$ for all h. This expression can be rewritten using a behavioral version of Roy's identity to get

$$\frac{V_I}{\lambda} + \gamma = 1 + \frac{B(h)}{M(h)} \tag{45}$$

where

$$\gamma = \frac{\delta(\theta;h)}{\lambda M(h)}\left(\frac{U_\theta}{V_I} - K_\theta\right) \tag{46}$$

is the bias-correction effect. The RHS is the same as in the standard model. The first term on the LHS, $\frac{V_I}{\lambda}$, is analogous to the standard model. However, in the model with optimizing agents I could express it in terms of marginal utilities of consumption through the agent's first order conditions. The new term γ captures the first order welfare effect of behavioral adjustments. The

[28] To save notation θ includes all choices, i.e. $\theta = \{C, \Pi\}$, and I suppress that the utility function might depend on the state history X.

expression $\frac{U_\theta}{V_I} - K_\theta$ in γ measures the welfare cost of the behavioral biases. Farhi and Gabaix (2015) call this term the "*behavioral wedge*" and Bernheim and Taubinsky (2018) refer to this as the "*price-metric measure of bias*". Intuitively, $\frac{U_\theta}{V_I} - K_\theta$ is the difference between the money-metric of marginal utilities, $\frac{U_\theta}{V_I}$, and the prices K_θ and therefore measures the degree of misoptimization. With rational optimizing agents this behavioral wedge is zero.[29] If agents fail to optimize their true preferences, γ is non-zero and is determined by the behavioral wedge multiplied with the reform induced change in behavior, $\delta(\theta; h)$. The sign of bias-correction term, γ, depends on the exact biases present and whether the reform corrects the sub-optimal behavior or amplifies the biases.

The two most prominent behavioral biases in the retirement context are myopia and statutory retirement ages as reference-points. In case of myopic behavior, individuals save too little, work not enough and retire too early.[30] Hence, a policy change that induces more savings, later retirement and so on has a positive bias correction term and reduces the direct welfare effect of taking a dollar away from retirees. If retirement ages serve as reference points, it is not clear what the true utility is or should be. The additional complication is that the planner can set the statutory retirement ages and might therefore be able to shift the reference point. How to evaluate welfare effects in this context is an open question and an interesting avenue for future research.

[29] Formally, a rational agent solves $\max_\theta U(\theta)$ subject to $K = \sum_{t=0}^{T-1} [1 + r_t]^{-t} E\left[y_t(x^t) - \tau(x^t) + b(x^t) - c_t(x^t) - q(p_t(x^t))\right] - I = 0$, which yields first order condition $U_\theta = V_I K_\theta$. Therefore, $\gamma = 0$.

[30] The consensus in the literature is to model myopic behavior with hyperbolic discounting and assume that the true utility function uses exponential discounting. However, Bernheim (2009, 2016) argues that it is not clear whether this is really a bias/mistake. It could also be interpreted as a present focus, i.e. living in the moment, that should not be corrected by the planner.

B Evaluation of the Early Retirement Age Reforms

Other Policy Variation 2000 Pension Reform: The 2000 pension reform also changed other margins but less dramatically than the ERA and these changes do not systematically vary by quarter of birth. The reform increased the penalty for claiming retirement benefits before the normal retirement age (NRA). Before the reform the pension coefficient is reduced by 2 percentage points for each year of claiming retirement benefits before the NRA. This penalty is capped at 10 percentage points. Moreover, the reduction can be at most 15 percent of the pension coefficient before the penalty.[31] The 2000 pension reform increases the penalty from 2 to 3 percentage points for each year of claiming before the NRA. The maximal penalty is slightly increased from 10 percentage points to 10.5 percentage points and the rule that the pension coefficient with penalty can at most be 15 percent lower than without penalty remains in place. However, the new penalty only applies to men born after September 1942 and to women born after September 1947. For men born before September 1942 and women born before September 1947, there was no change in the pension formula if they claimed at the earliest possible age. If they claimed at older ages the change in the penalty from 2 to 3 percentage points is phased in by quarter of birth. This phase-in led to small differences in the replacement rate between quarter of births for claiming after the ERA. However, the differences in the replacement rate between two adjacent quarters of birth are around 0.2 percentage points (this corresponds to a difference in pension of 0.25%). The treatment groups of the ERA increase of the 2000 pension reform are men born between October 1940 and September 1942 and women born born between October 1945 and September 1947. Therefore, my treatment groups are subject to the transition rules and are not strongly treated by the increase in the penalty. Especially, the differences between two adjacent quarters of birth are minimal. The 2000 pension reform also temporarily extended the maximum duration of unemployment benefits from 1 to 1.5 years for a subgroup of individuals. To qualify for the benefit extension individuals need at least 15 years of employment in the past 25 years and only certain birth cohorts in certain years were eligible. In 2000, only men born in 1940 and women born in 1945 were eligible. In 2001, men born in 1940-41 and women born in 1945-46 were eligible and in 2002 men born in 1940-1942 and women born in 1945-47 were eligible. The UI benefit extension ended in December 2002.

[31] That is, the pension of early claimaints can at most be 15 percent lower compared to their pension if they retired at NRA (with the same earnings history and number of insurance years).

B.1 Results for the Pension Reform in 2000

B.2 Difference-in-Difference Estimation for Adjacent Birth Quarters

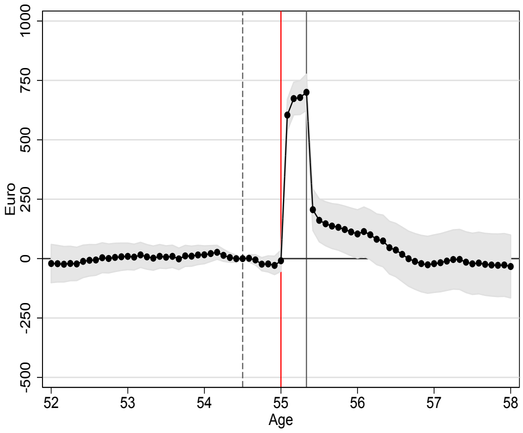
An alternative to my approach in the main text is to only compare adjacent quarters of birth with a difference-in-difference approach. In this case the counterfactual is different. For each group I estimate the effect of increasing the ERA by two months starting at the control group's ERA. The potential advantage of this approach is that adjacent quarters of birth might be more comparable and I observe them at almost the same point in time at a given age. The drawback is that sample size becomes relatively small. I implement this comparison with the following regression

$$Y_{it} = \alpha + \sum_{k=ERA-6}^{ERA+6} \beta_k * Treat_i * I[age_{it} = k] + \sum_{k=ERA-6}^{ERA+6} \kappa_k * I[age_{it} = k] + Treat_i + \lambda_t + \varepsilon_{it}. \quad (47)$$

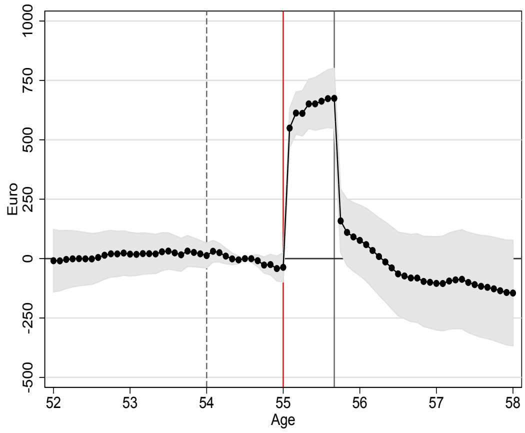
To make all changes directly comparable I normalize ages relative to the control groups ERA. Figures 29 to 32 plot the β_k estimates of the fiscal revenue effect as well as the mechanical effect of the ERA increase. We see the same patterns as in the analysis of the main text. The fiscal revenue effect is predominantly driven by the mechanical effect and in some cases the multiplier is below one. Table presents 5 the detailed numbers. Due to the much lower sample size, the estimates become a bit more jumpy but the multipliers are still centered around one.

Figure 18: DiD Estimates of Fiscal Revenue by Age: Women Reform 2000

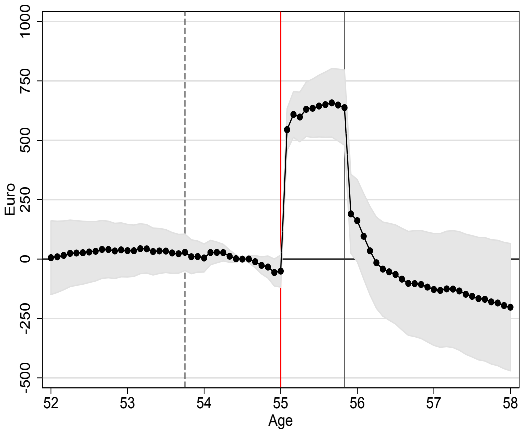
(a) ERA 55 + 4 months



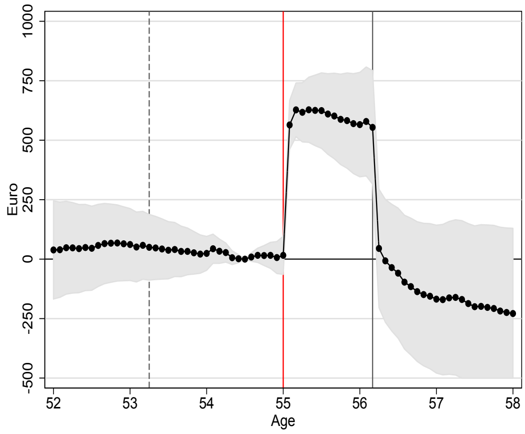
(b) ERA 55 + 8 months



(c) ERA 55 + 10 months



(d) ERA 56 + 2 months



(e) ERA 56 + 4 months

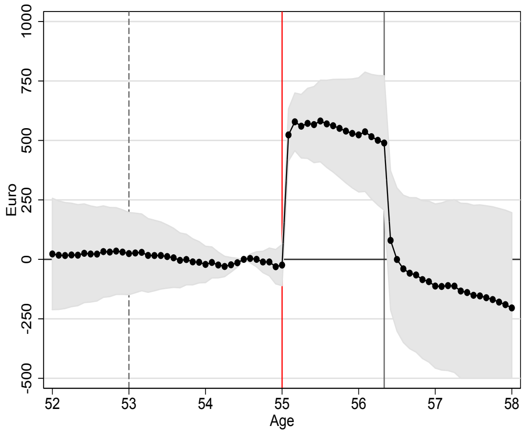
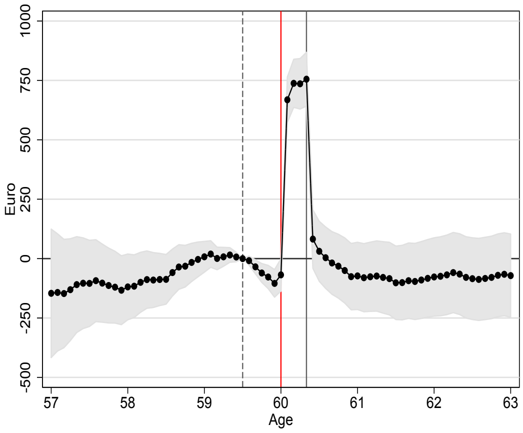
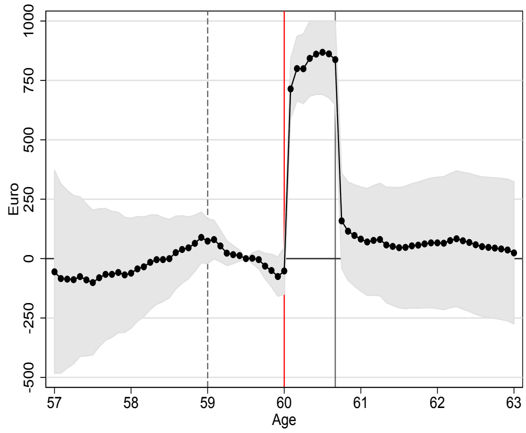


Figure 19: DiD Estimates of Fiscal Revenue by Age: Men Reform 2000

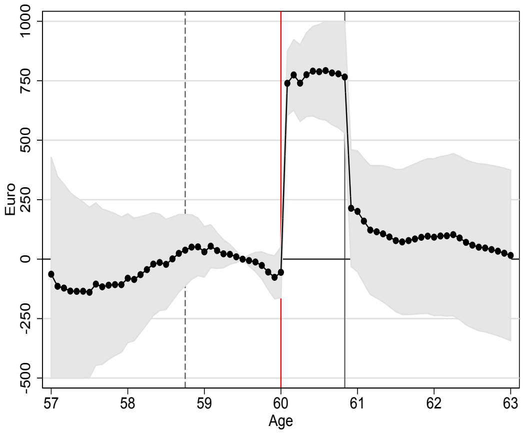
(a) ERA 60 + 4 months



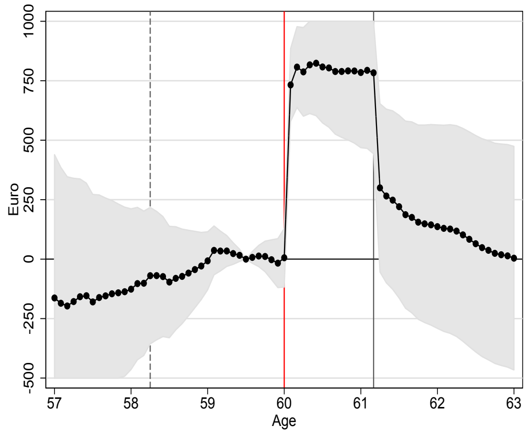
(b) ERA 60 + 8 months



(c) ERA 60 + 10 months



(d) ERA 61 + 2 months



(e) ERA 61 + 4 months

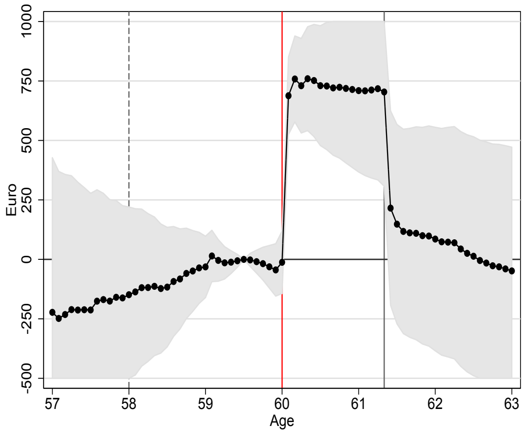


Figure 20: Mechanical Effect for Women, Reform in 2000

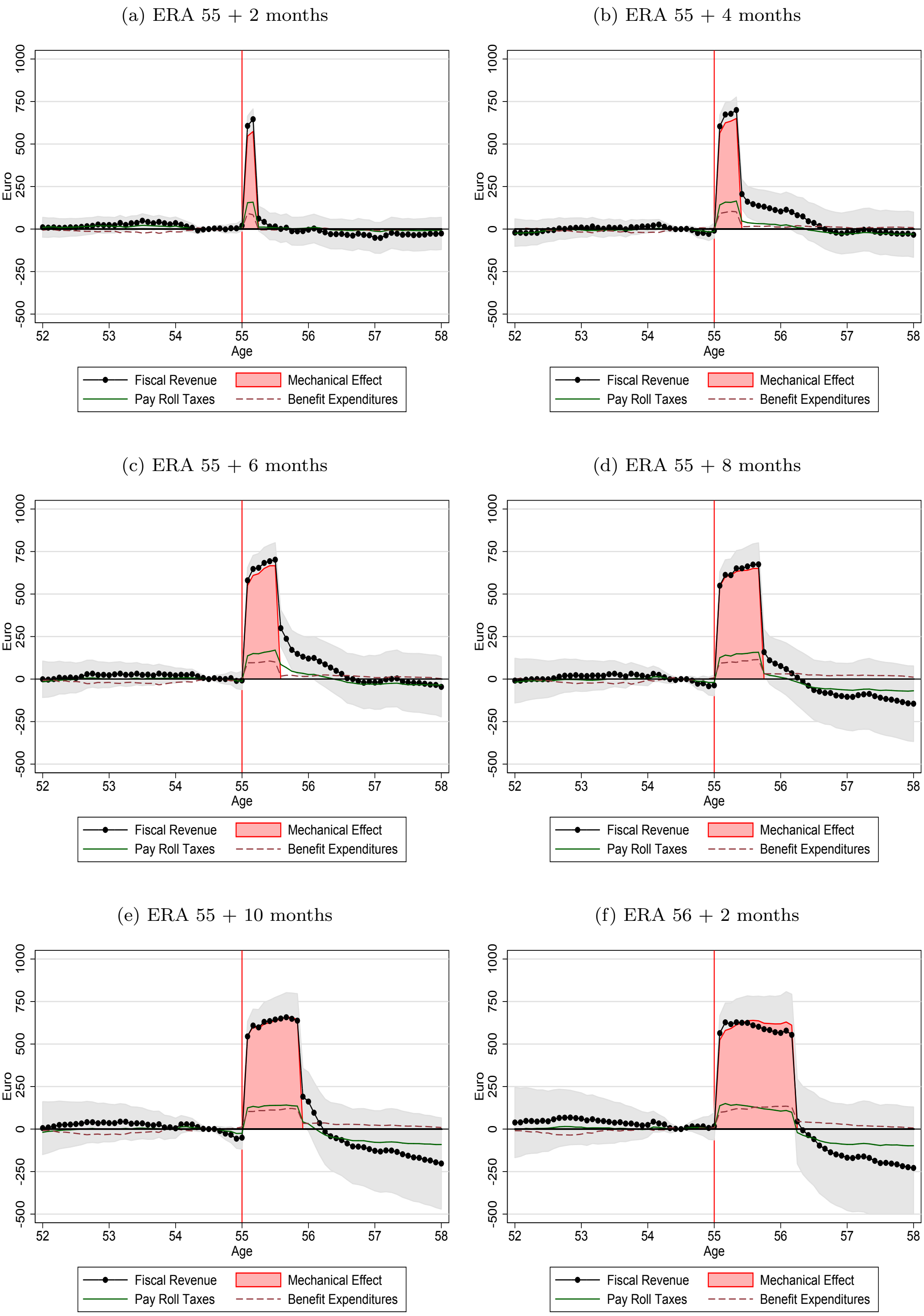
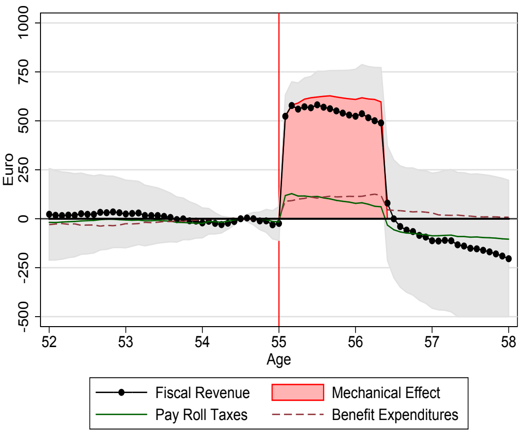


Figure 21: Mechanical Effect for Women, Reform in 2000

(a) ERA 56 + 4 months



(b) ERA 56 + 6 months

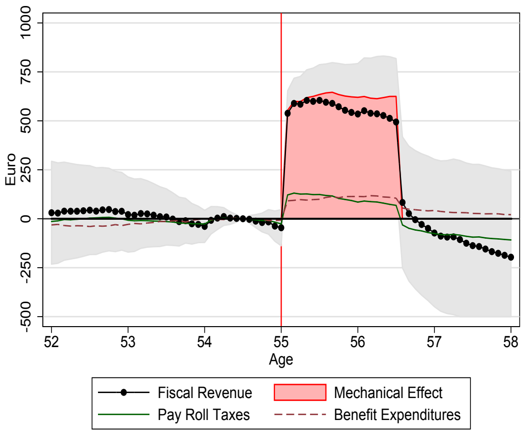
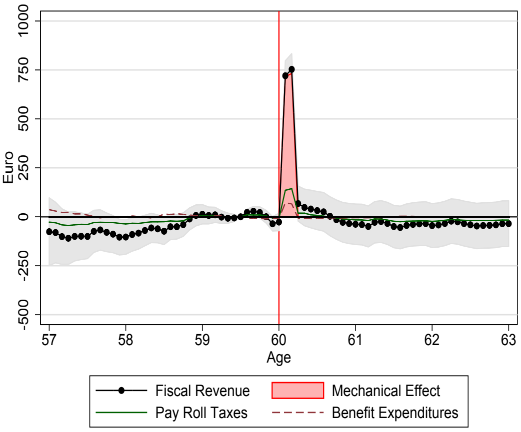


Table 5: Multipliers for the ERA Reform in 2000

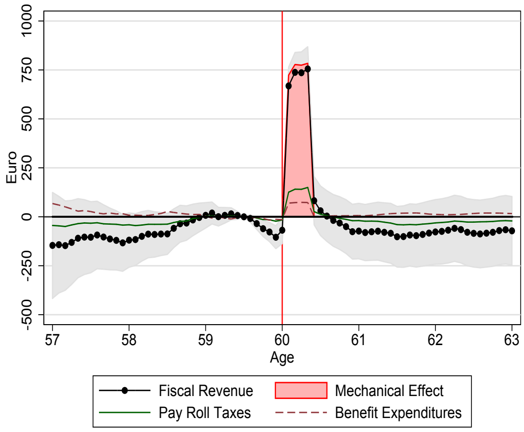
ERA	Total Fiscal Effect	Mechanical	Behavioral	Multiplier (1+B/M)
Women				
ERA 55 + 2 months	1038	1120	-83	0.93
ERA 55 + 4 months	1227	1216	11	1.01
ERA 55 + 6 months	887	984	-97	0.90
ERA 55 + 8 months	972	823	148	1.18
ERA 55 + 10 months	1008	985	23	1.02
ERA 56	1029	892	138	1.15
ERA 56 + 2 months	934	885	50	1.06
ERA 56 + 4 months	1003	970	33	1.03
ERA 56 + 6 months	757	872	-115	0.87
Men				
ERA 60 + 2 months	1249	1448	-199	0.86
ERA 60 + 4 months	1528	1400	128	1.09
ERA 60 + 6 months	1381	1373	8	1.01
ERA 60 + 8 months	1103	1145	-42	0.96
ERA 60 + 10 months	1347	1213	134	1.11
ERA 61	1148	867	281	1.32
ERA 61 + 2 months	1052	1264	-212	0.83
ERA 61 + 4 months	1174	822	352	1.43
ERA 61 + 6 months	955	845	110	1.13

Figure 22: Mechanical Effect for Men, Reform in 2000

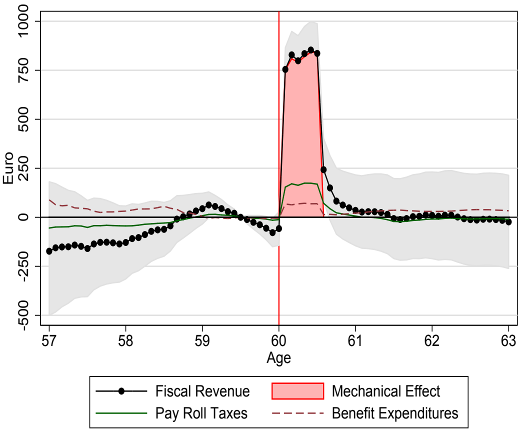
(a) ERA 60 + 2 months



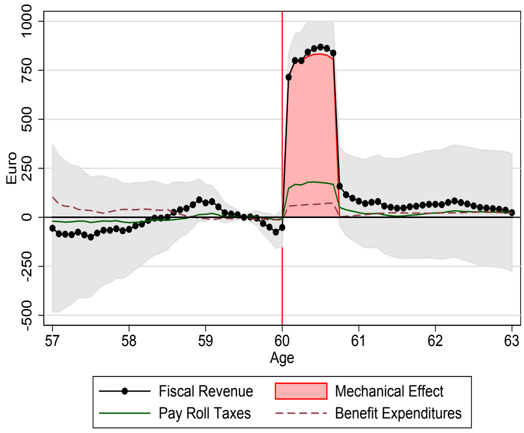
(b) ERA 60 + 4 months



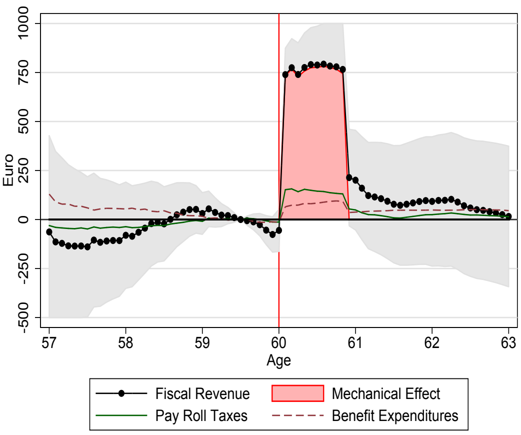
(c) ERA 60 + 6 months



(d) ERA 60 + 8 months



(e) ERA 60 + 10 months



(f) ERA 61 + 2 months

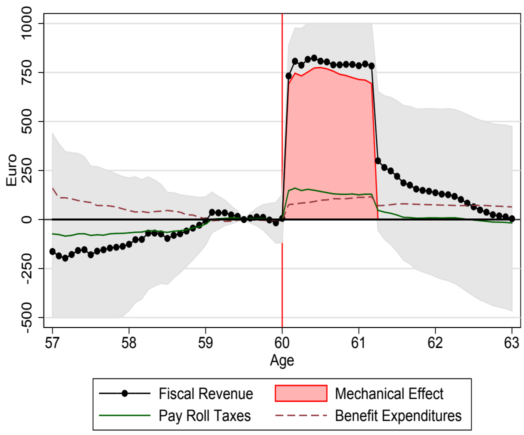


Figure 23: Mechanical Effect for Men, Reform in 2000

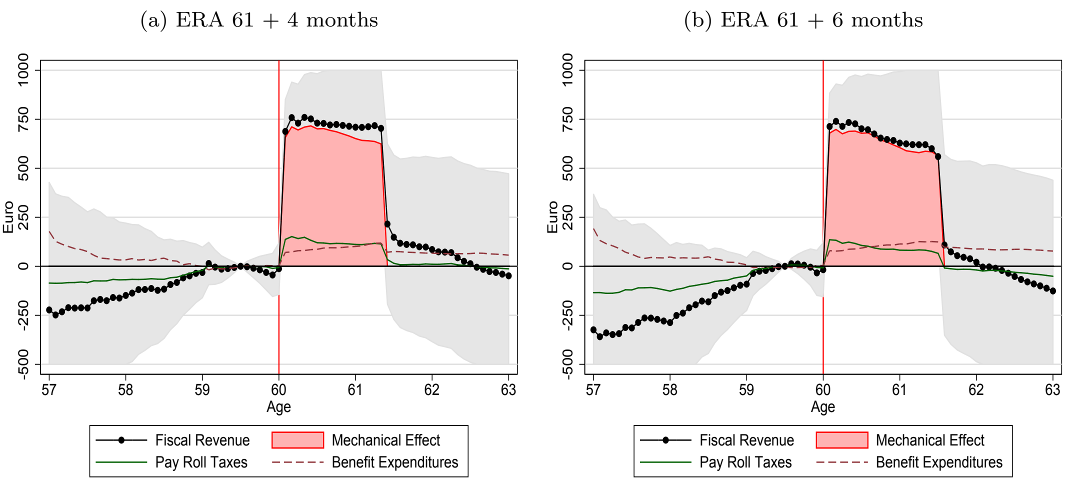
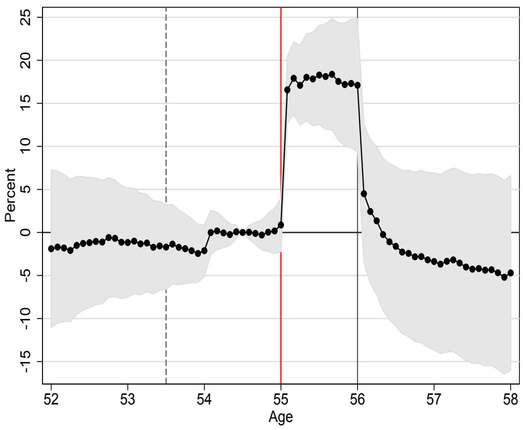
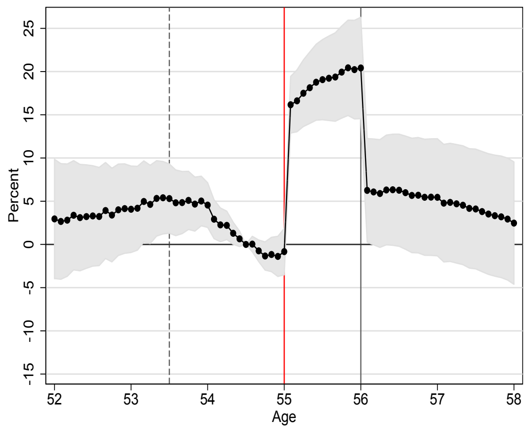


Figure 24: DiD Estimates for Labor Market Status by Age: Women ERA 56 months

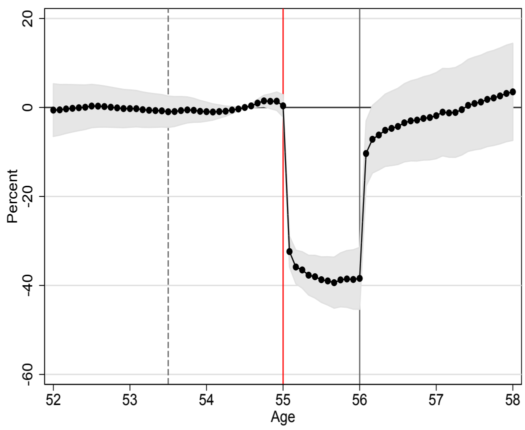
(a) Employment



(b) Unemployment



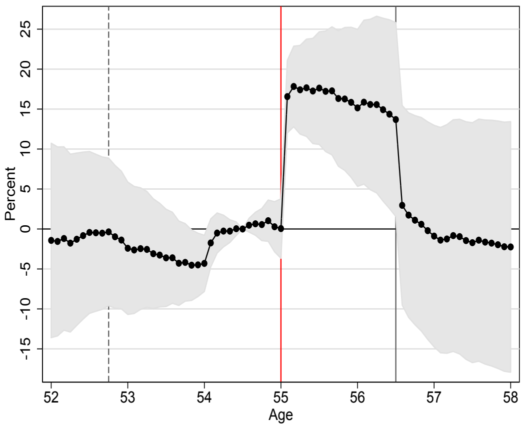
(c) Retirement (OA and DI)



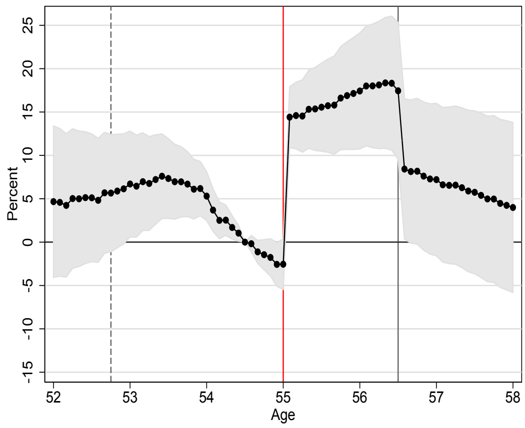
Notes: This figure shows the DiD estimates for employment, unemployment and retirement.

Figure 25: DiD Estimates for Labor Market Status by Age: Women ERA 56.5 months

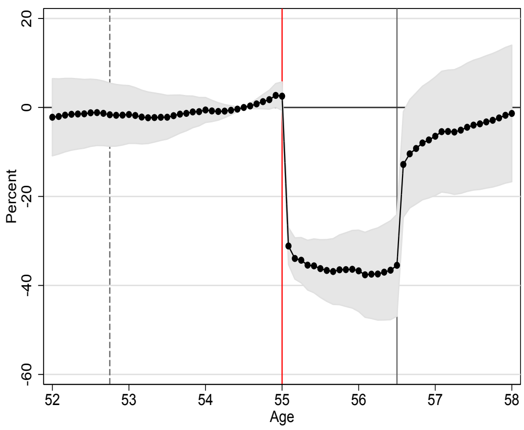
(a) Employment



(b) Unemployment

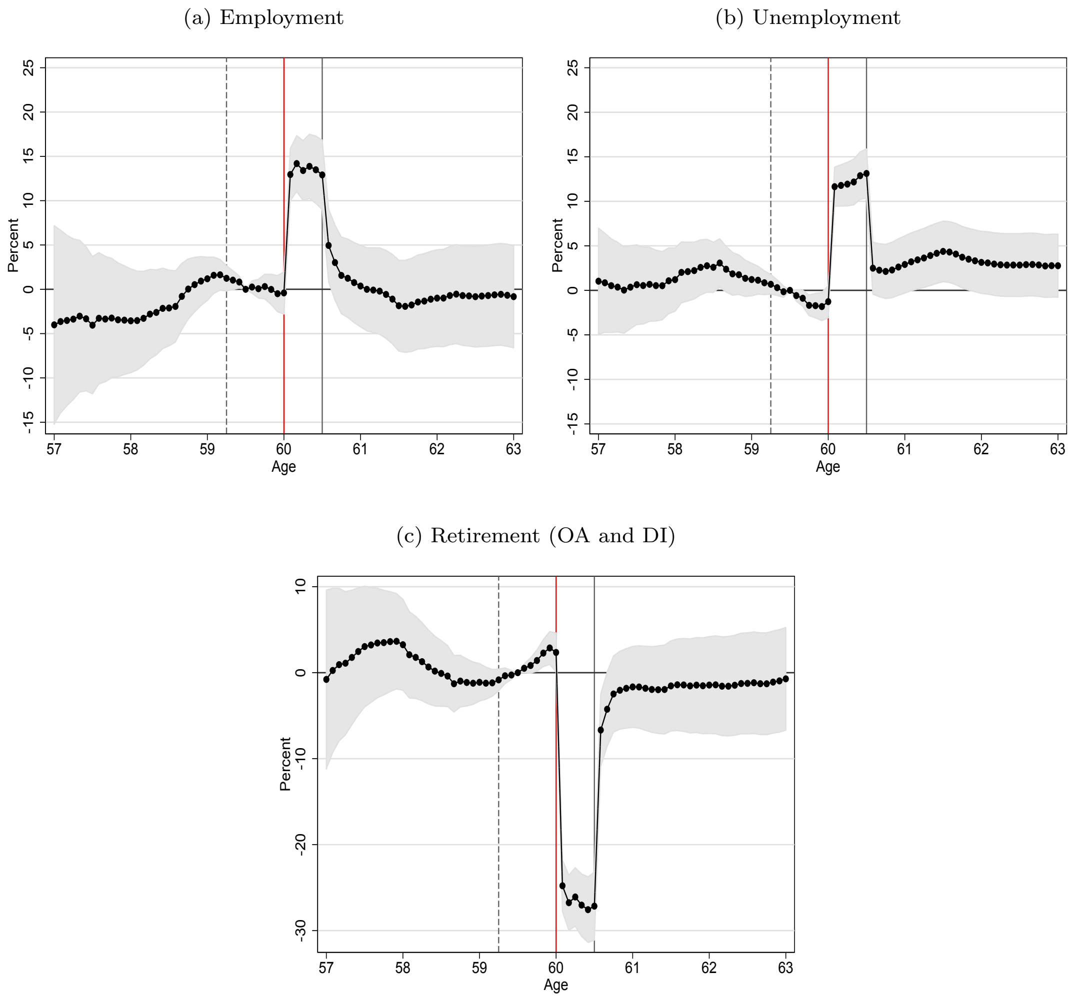


(c) Retirement (OA and DI)



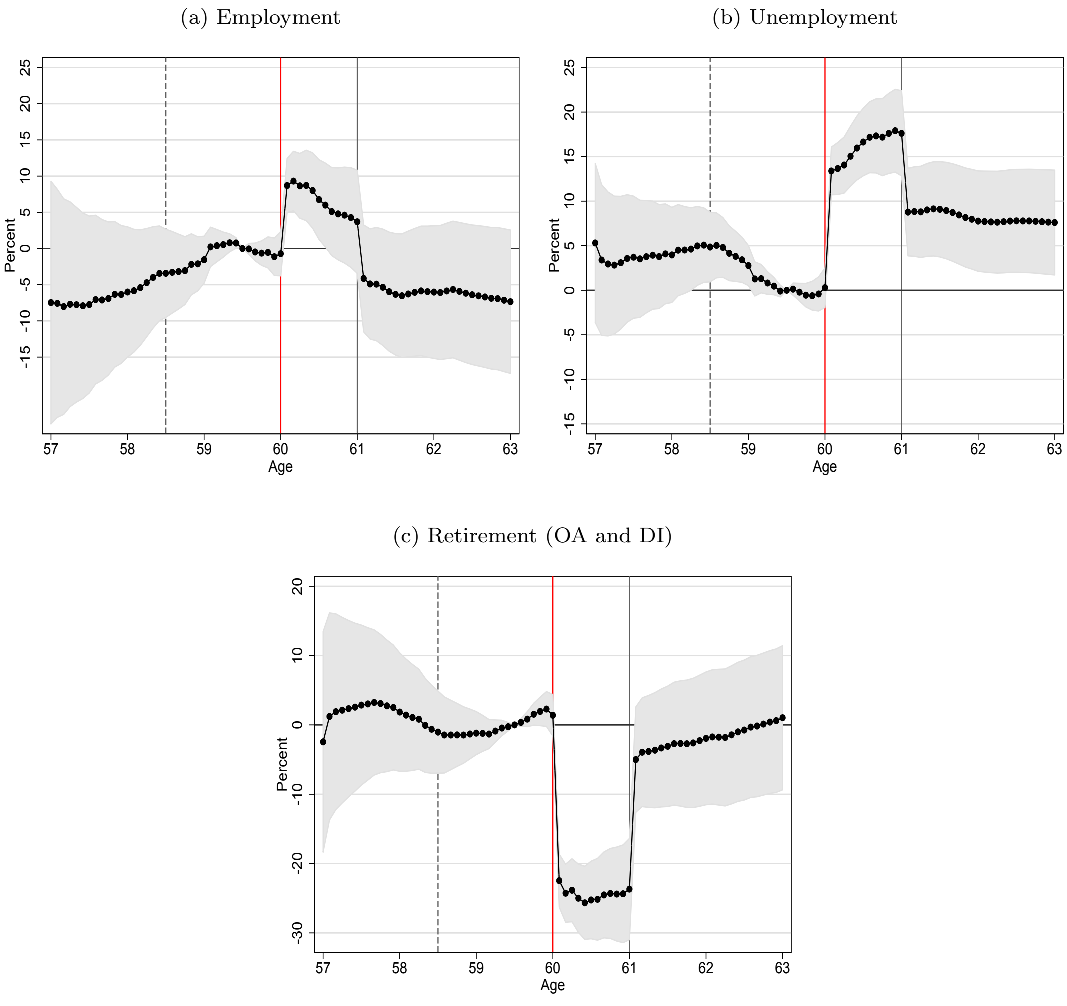
Notes: This figure shows the DiD estimates for employment, unemployment and retirement.

Figure 26: DiD Estimates for Labor Market Status by Age: Men ERA 60 + 6 months



Notes: This figure shows the DiD estimates for employment, unemployment and retirement.

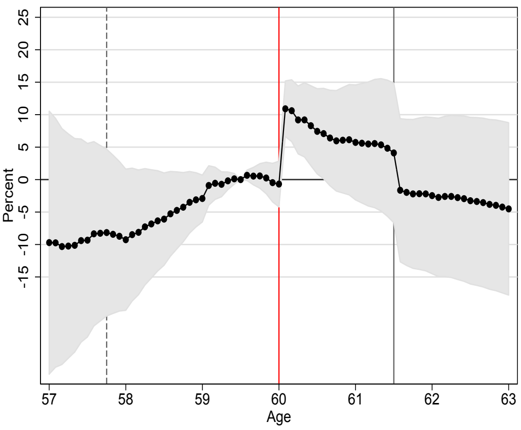
Figure 27: DiD Estimates for Labor Market Status by Age: Men ERA 61



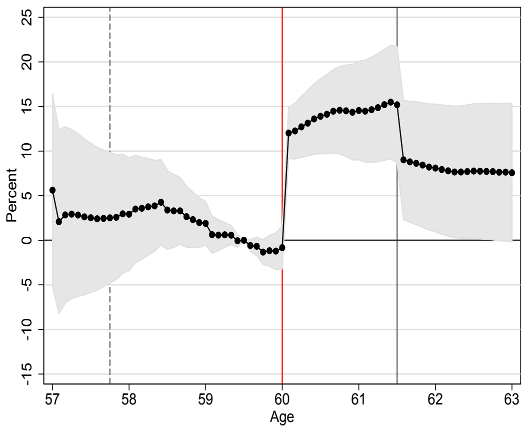
Notes: This figure shows the DiD estimates for employment, unemployment and retirement.

Figure 28: DiD Estimates for Labor Market Status by Age: Men ERA 61 + 6 months

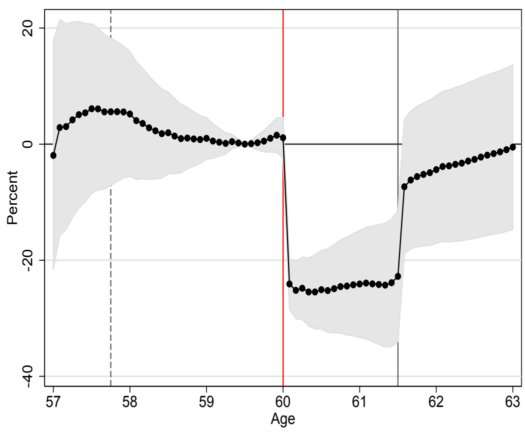
(a) Employment



(b) Unemployment



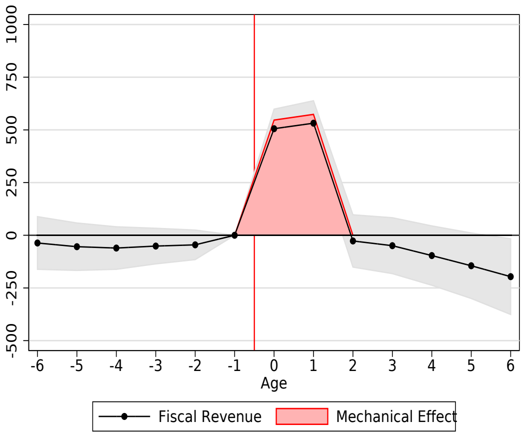
(c) Retirement (OA and DI)



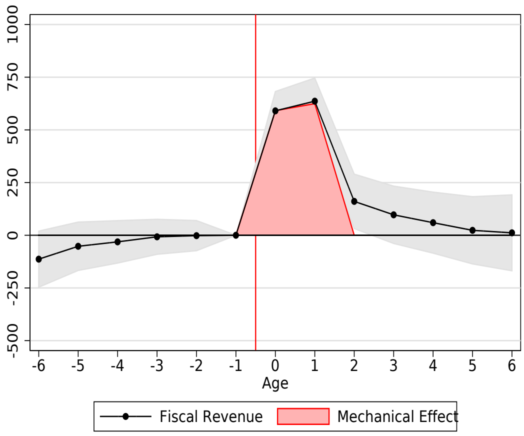
Notes: This figure shows the DiD estimates for employment, unemployment and retirement.

Figure 29: Mechanical Effect Women Reform 2000

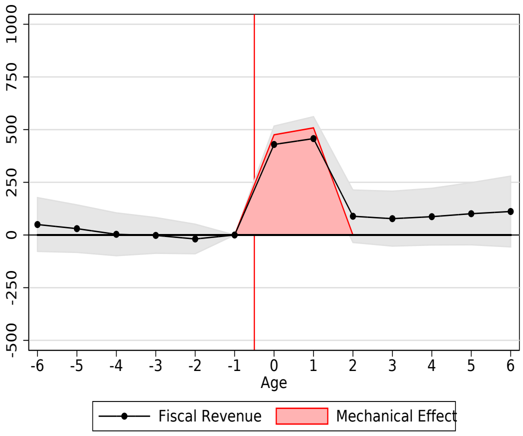
(a) ERA 55 + 2 months



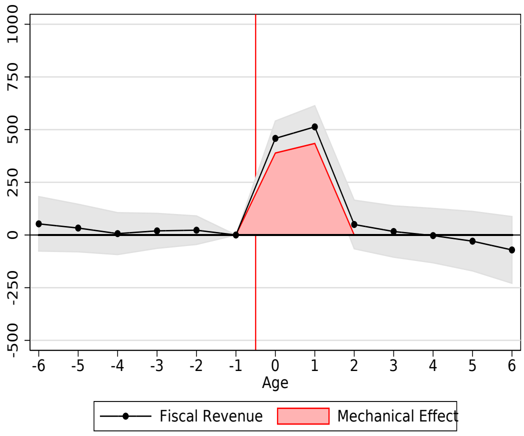
(b) ERA 55 + 4 months



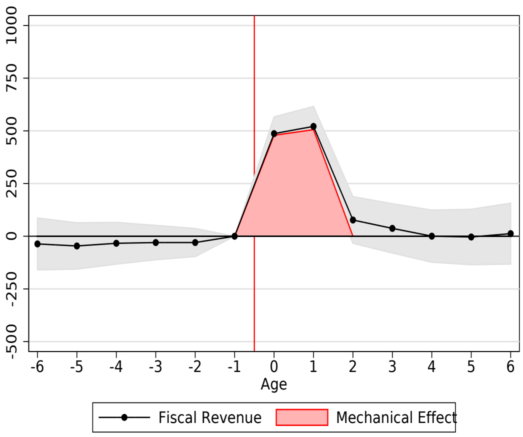
(c) ERA 55 + 6 months



(d) ERA 55 + 8 months



(e) ERA 55 + 10 months



(f) ERA 56 months

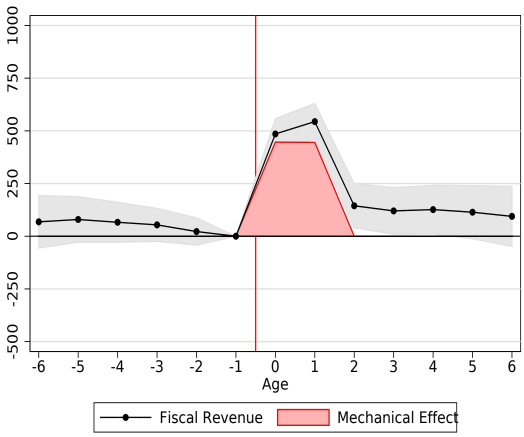
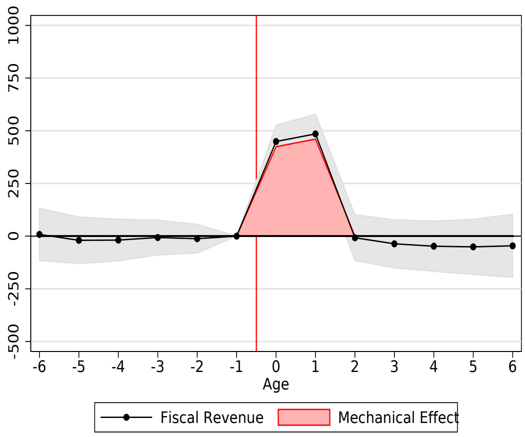
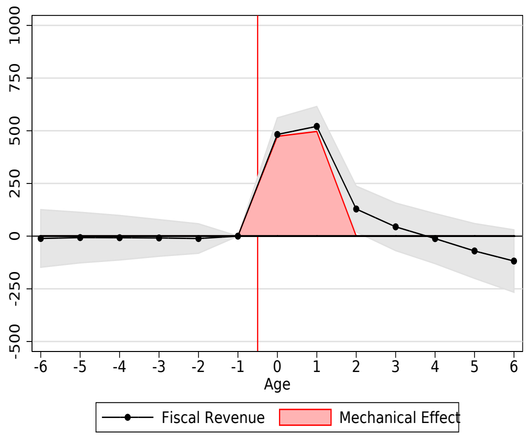


Figure 30: Mechanical Effect Women Reform 2000

(a) ERA 56 + 2 months



(b) ERA 56 + 4 months



(c) ERA 56 + 6 months

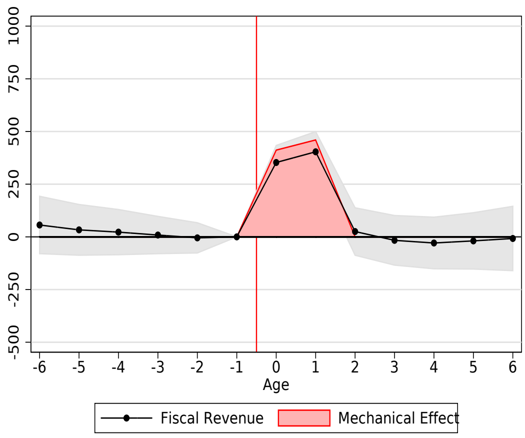
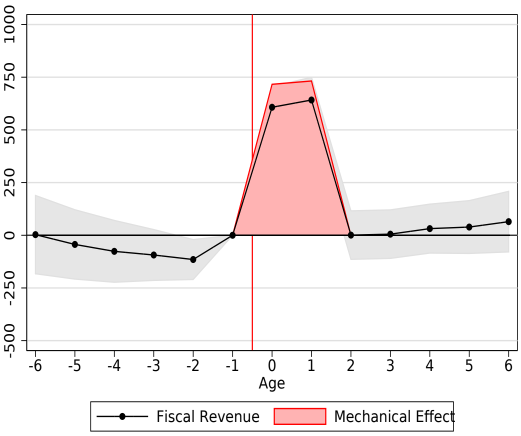
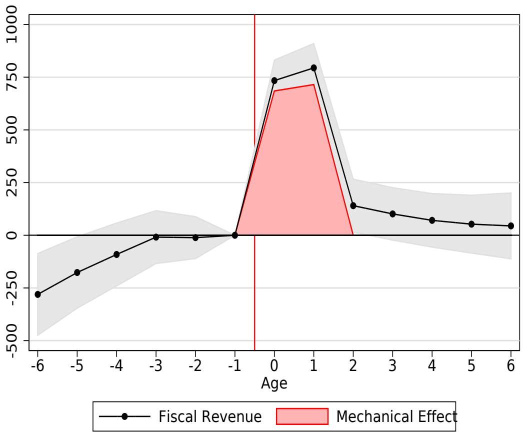


Figure 31: Mechanical Effect Men Reform 2000

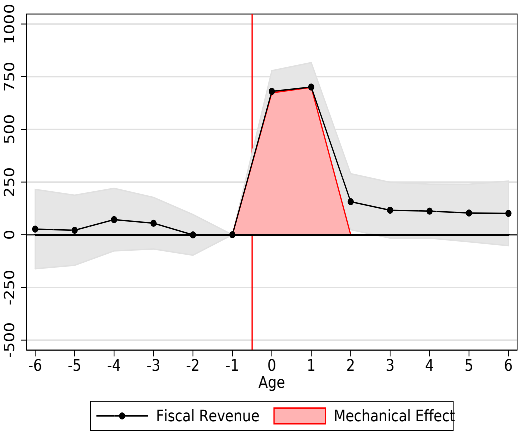
(a) ERA 60 + 2 months



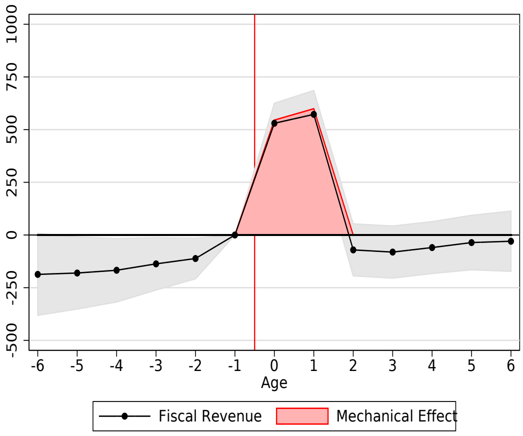
(b) ERA 60 + 4 months



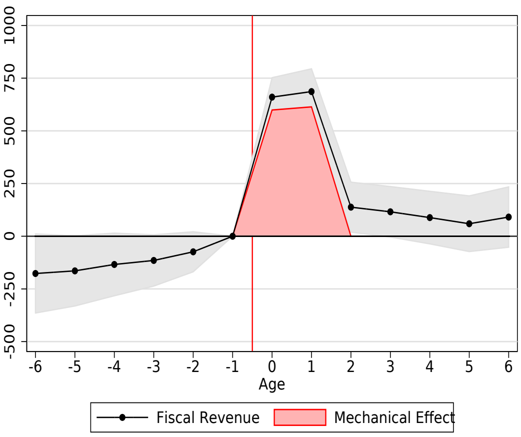
(c) ERA 60 + 6 months



(d) ERA 60 + 8 months



(e) ERA 60 + 10 months



(f) ERA 61 months

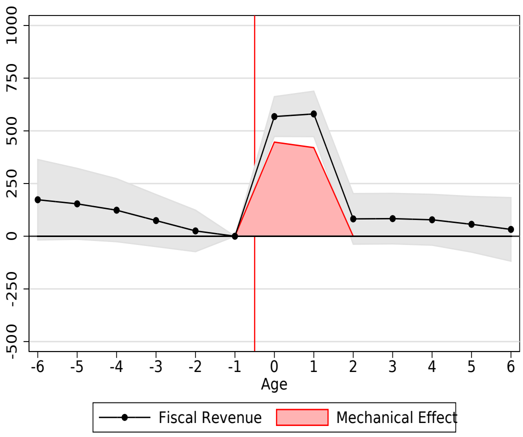
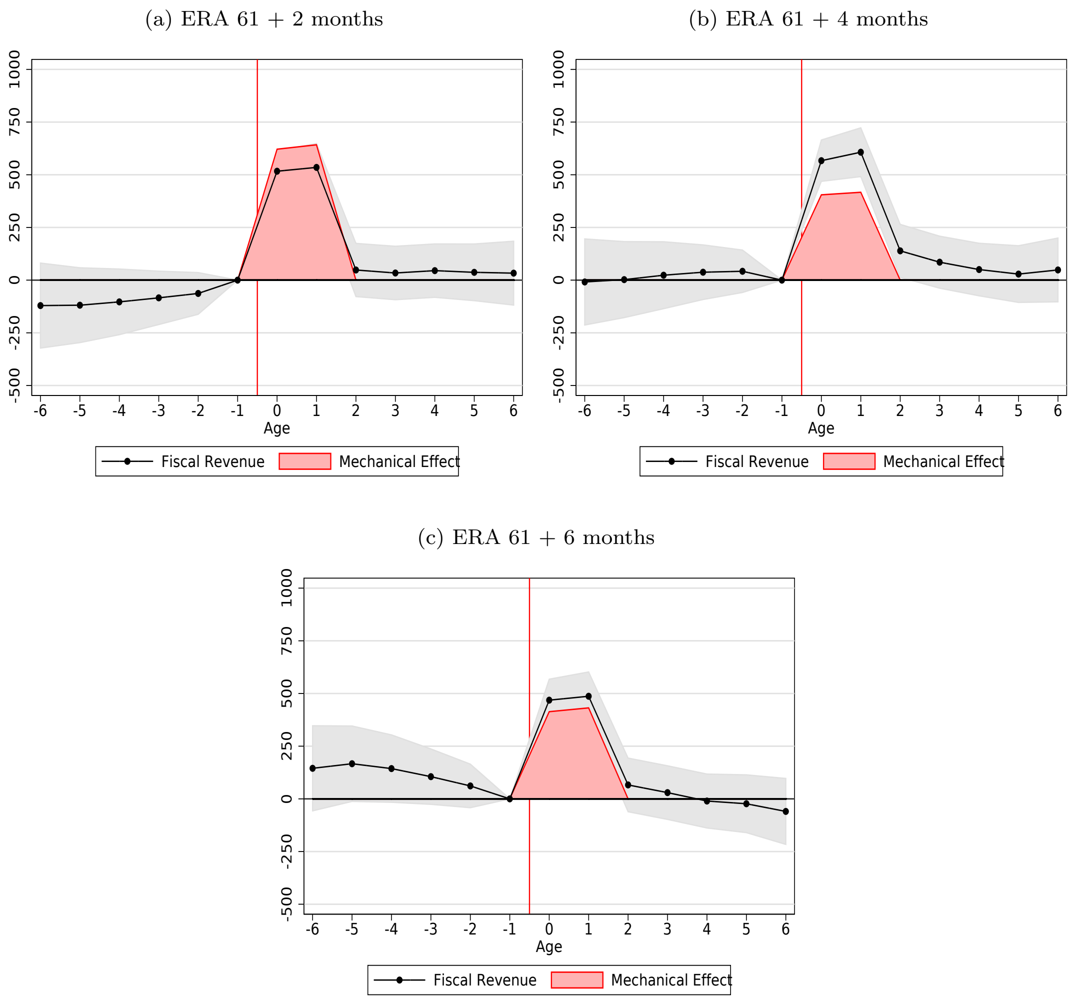


Figure 32: Mechanical Effect Men Reform 2000



C Pension Reform in 2003

The 2003 pension reform further increased the ERA from 56.5 to 60 for women and from 61.5 to 65 for men. However, the reform also changed the penalty for retiring early significantly and the replacement rate for a given number of insurance years. This makes it more difficult to separate the ERA from other changes. Nevertheless, the results are consistent with the results from the 2000 pension reform as multipliers are still around 1.

Other Policy Variation 2003 Pension Reform: The 2003 pension reform reduces benefit generosity by reducing the pension coefficient for a given number of insurance years. Before the reform, each insurance year increased the pension coefficient by 2 percentage points, after the reform this is reduced to 1.78 percentage points. This change is also phased-in by claiming year: 1.96 pp in 2004, 1.92 pp in 2005, 1.88 pp in 2006, 1.84 pp in 2007, 1.80 pp in 2008, 1.78 pp since 2009. The penalty for claiming before the NRA is increased from 3 percentage points per year to 4.2 percentage points without transition rules. Furthermore, the assessment period changes from the best 15 to the best 40 years. This change is phased-in: Since 2004 the assessment period increases by one year each year until the assessment period is 40 years in 2028. The 2004 pension reform reintroduces the possibility of early retirement by creating the so-called corridor pension. Individuals with at least 37.5 insurance years. Until 2032 pension cuts resulting from the pension reform were capped at 10%, except for the losses due to the increase in the ERA. However, the 2004 pension reform sets this cap at 5% until 2024 and then gradually increases the cap to 10%.

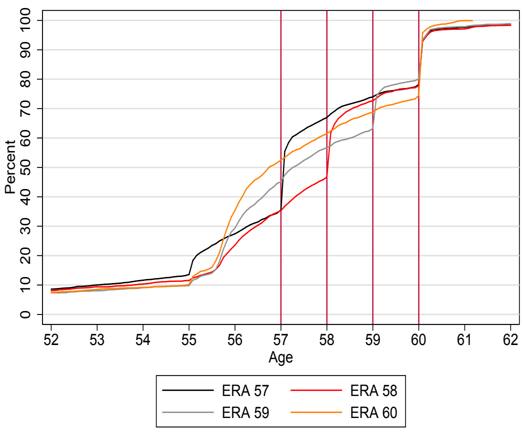
C.1 Empirical Strategy for the Pension Reform in 2003

I exploit the variation in the ERA by quarter of birth in a cohort difference-in-difference specification analogous to the specification in the main text. I focus here on women because for men the ERA increase beyond age 62 was irrelevant for most individuals (see figure 34). Individuals with more than 37.5 insurance years can still retire at age 62 through the so called "corridor pension". Hence, most men were not affected by the ERA increase.

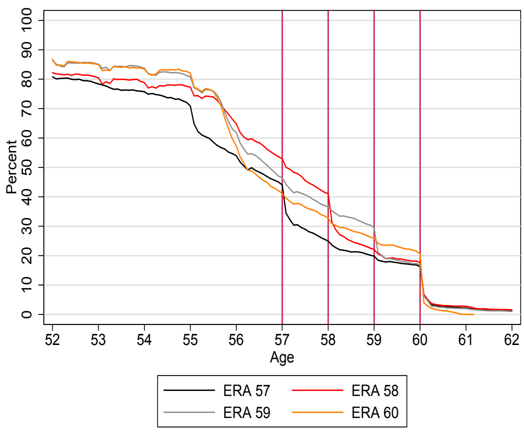
The control group for women is women born in 1948 Q2, with ERA 56 and 10 months. I take this control group since they are already affected by the change in the penalty and after this group the ERA increases by 1 month for each quarter of birth.

Figure 33: Reform 2003: Women Labor Market Status by Age

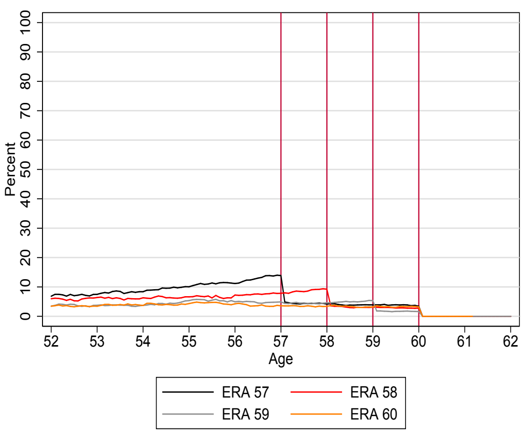
(a) Retirement (Disability and Old-Age)



(b) Employment



(c) Unemployment



(d) Other

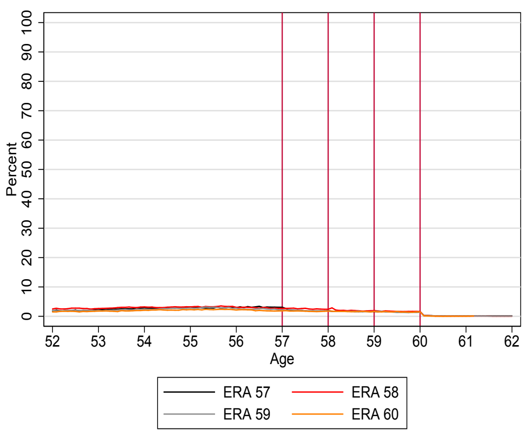
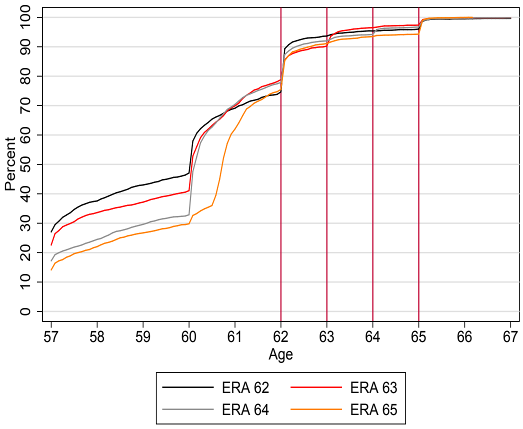
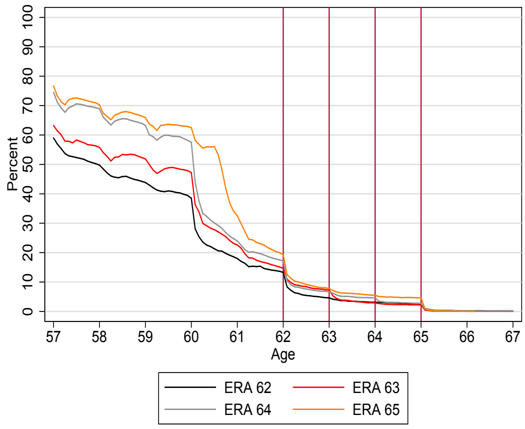


Figure 34: Reform 2003: Men Labor Market Status by Age

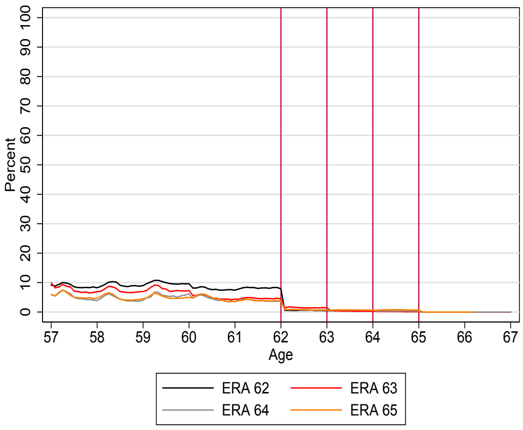
(a) Retirement (Disability and Old-Age)



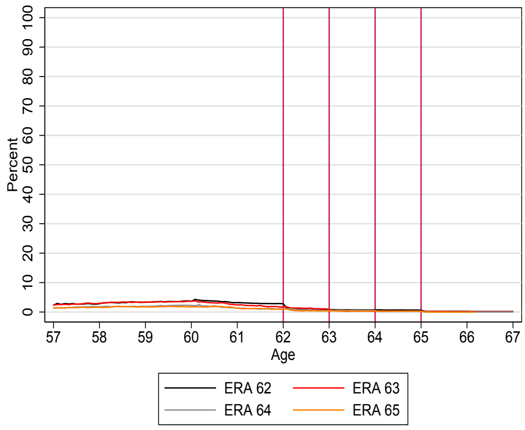
(b) Employment



(c) Unemployment



(d) Other



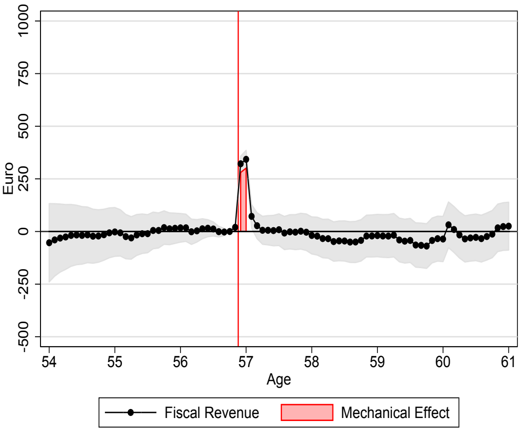
C.2 Descriptive Figures for the Pension Reform in 2003

C.3 Results Pension Reform 2003

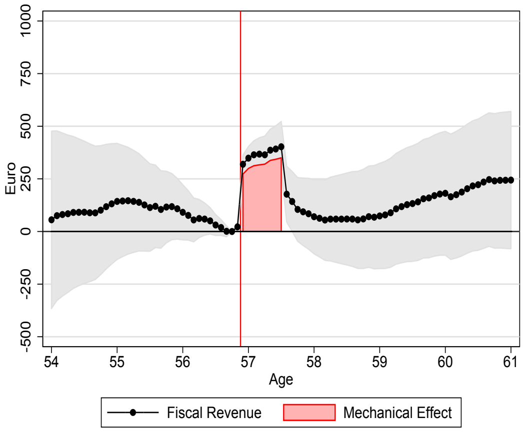
For brevity, I focus here on fiscal outcomes and the multiplier. Figure 35 shows the fiscal revenue and mechanical effect for women for selected treatment groups. Table shows the multiplier for all groups for women. The main take-away from these tables is that the multiplier is still relatively small and around 1. However, the estimates of the multiplier can be a mix between increasing ERA, level shifts in generosity and changes in actuarial fairness (because the 2003 reform significantly changed other margins of the pension formula as discussed at the beginning).

Figure 35: DiD Estimates Fiscal Revenue by Age: Men Reform 2003

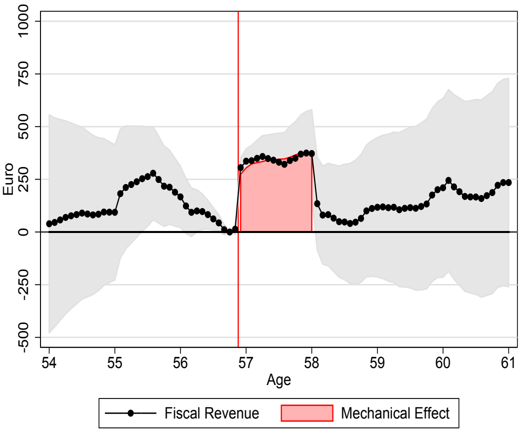
(a) ERA 57



(b) ERA 57 + 6 months

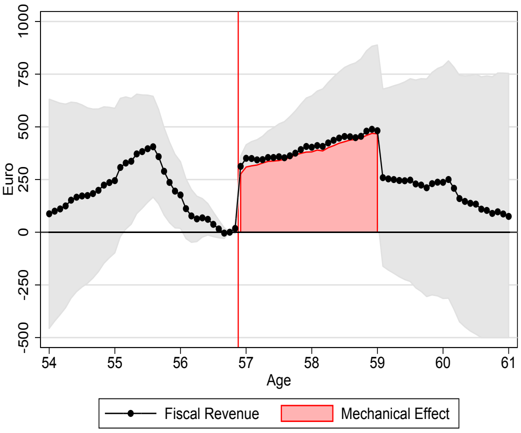


(c) ERA 58



(d) ERA 58 + 6 months

(e) ERA 59



(f) ERA 59 + 6 months

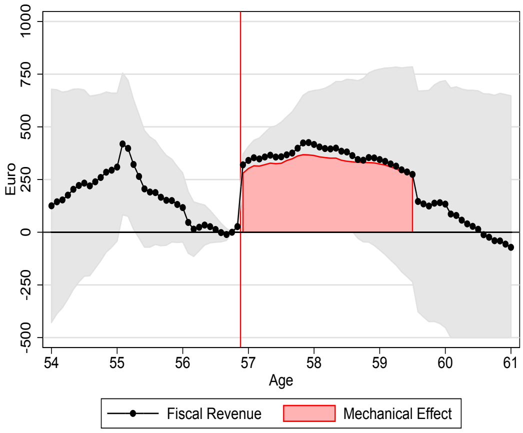


Table 6: Multiplier ERA Reform 2003

ERA	Fiscal Revenue Effect	Mechanical	Behavioral	Multiplier
Women				
ERA 56 + 11 months	303	260	43	1.16
ERA 57 months	665	582	83	1.14
ERA 57 + 1 months	1027	899	129	1.14
ERA 57 + 2 months	1328	1217	110	1.09
ERA 57 + 3 months	1719	1536	183	1.12
ERA 57 + 4 months	2083	1853	230	1.12
ERA 57 + 5 months	2364	2075	289	1.14
ERA 57 + 6 months	2945	2552	393	1.15
ERA 57 + 7 months	3291	2910	380	1.13
ERA 57 + 8 months	3566	3282	284	1.09
ERA 57 + 9 months	3752	3389	363	1.11
ERA 57 + 10 months	4373	3928	446	1.11
ERA 57 + 11 months	4725	4379	346	1.08
ERA 58 months	4835	4789	46	1.01
ERA 58 + 1 months	5299	5179	120	1.02
ERA 58 + 2 months	5944	5751	193	1.03
ERA 58 + 3 months	7058	6563	495	1.08
ERA 58 + 4 months	6526	6280	247	1.04
ERA 58 + 5 months	6908	6633	275	1.04
ERA 58 + 6 months	7897	7423	474	1.06
ERA 58 + 7 months	8894	8331	564	1.07
ERA 58 + 8 months	8936	8378	558	1.07
ERA 58 + 9 months	9378	8841	537	1.06
ERA 58 + 10 months	10305	9473	832	1.09
ERA 58 + 11 months	11545	10536	1009	1.10
ERA 59 months	10450	9928	522	1.05
ERA 59 + 1 months	11552	10949	603	1.06
ERA 59 + 2 months	12754	11911	843	1.07
ERA 59 + 3 months	13514	12453	1061	1.09
ERA 59 + 4 months	7386	10533	-3147	0.70
ERA 59 + 5 months	7795	10602	-2807	0.74
ERA 59 + 6 months	7933	10506	-2573	0.76
ERA 59 + 7 months	5607	9514	-3907	0.59
ERA 59 + 8 months	4735	8671	-3936	0.55
ERA 59 + 9 months	4227	7626	-3399	0.55
ERA 59 + 10 months	6251	8470	-2218	0.74

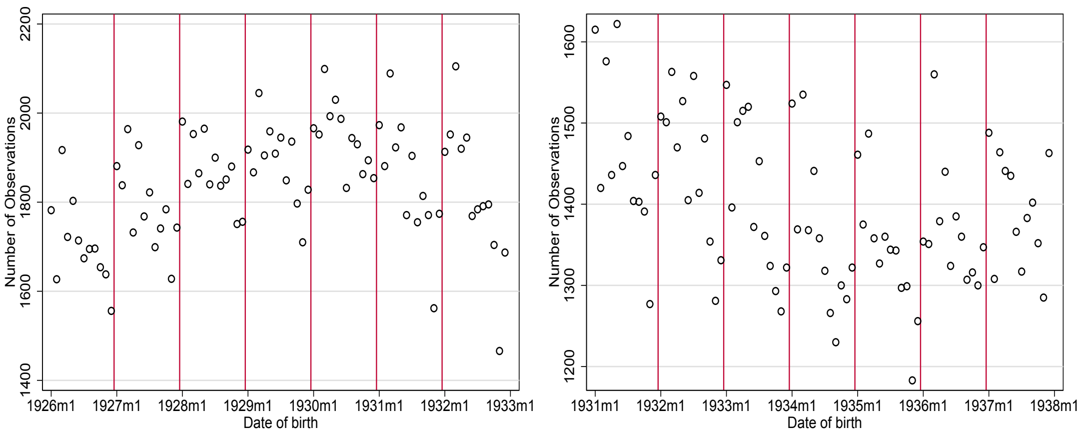
D Evaluation of Pension Level Reform

D.1 Balance Checks

Figure 36 shows that the number of observations around January is not balanced. This is potentially a problem. However, I do not find any effects for labor market outcomes or fiscal revenue before the reform. Moreover, figure 14 in the main text shows that the effects are driven by the treated group. I do not find effects for the placebo group that is unaffected by the reform. This speaks for the validity of my RDD despite the unbalanced number of observations around the cut-off.

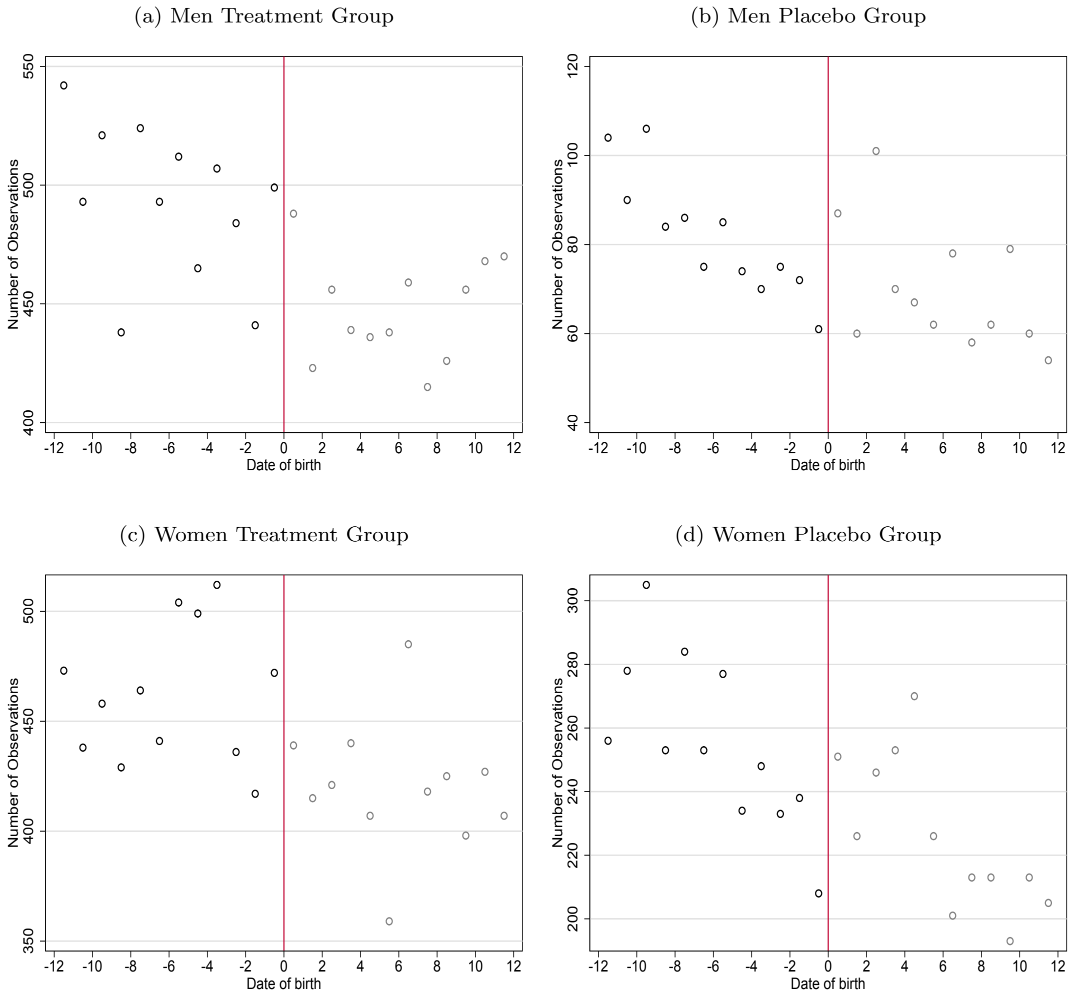
Figure 36: Number of Observations by Date of Birth

(a) Men (b) Women



Notes: This figure shows the number of observations by date of birth. Around January the number of observations is not balanced. However, this is unlikely to be driven by manipulation as a response to the reform.

Figure 37: Number of Observations by Date of Birth

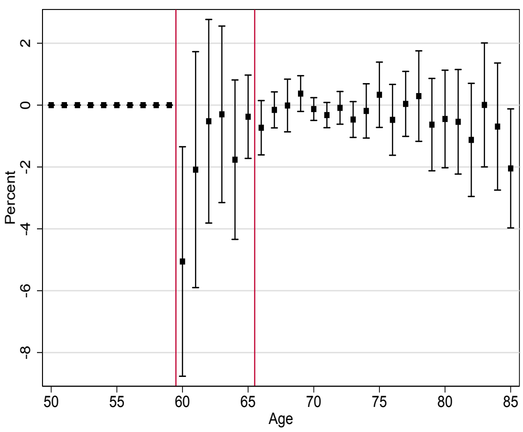


Notes: This figure shows the number of observations by date of birth for the treatment and placebo groups. The cutoff is at 1.1.1928 for men and 1.1.1933 for women.

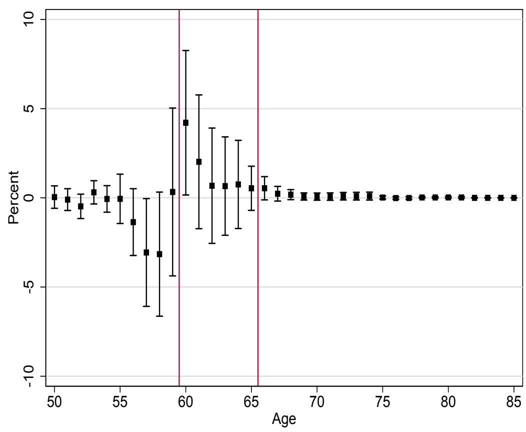
D.2 Additional Figures and Tables for Men

Figure 38: RD Estimates by Age: Labor Supply Men Treatment and Placebo

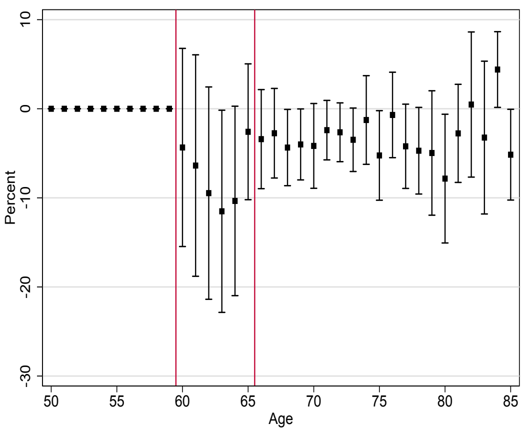
(a) Treatment group: Retirement (OA and DI)



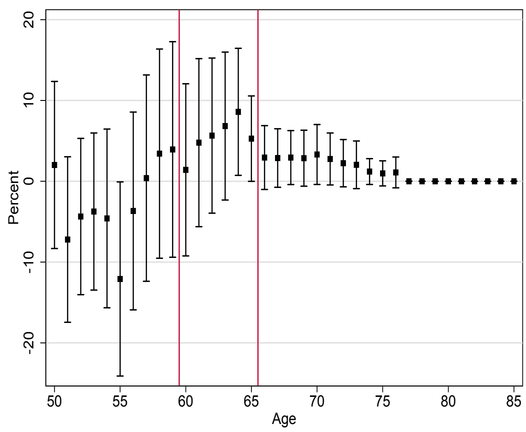
(b) Treatment group: Employment



(c) Placebo group: Retirement (OA and DI)

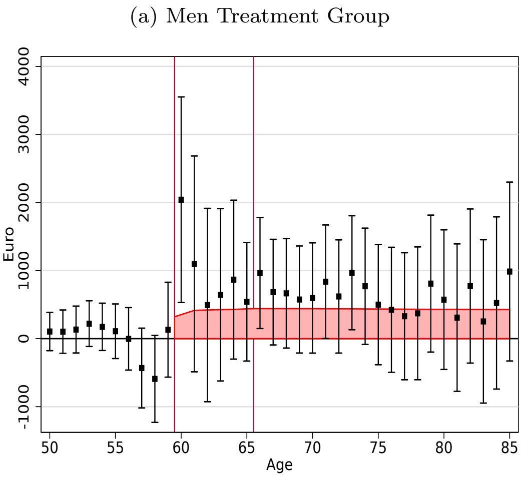


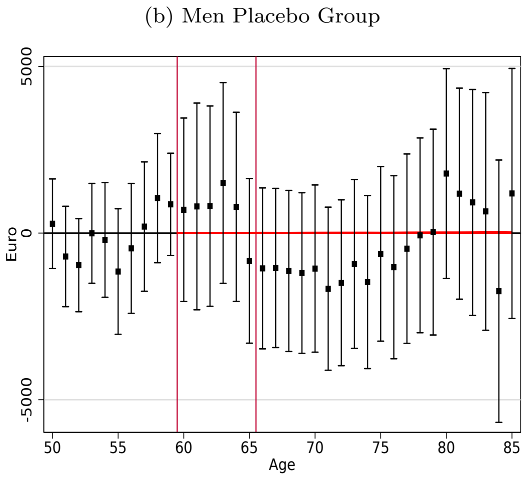
(d) Placebo group: Employment



Notes: This figure plots the RD estimates at each age between 50 and 85 based on local linear regressions with a bandwith of 8 months.

Figure 39: RD Estimates by Age: Fiscal Effect Men Treatment and Placebo





Notes: This figure plots the RD estimates for fiscal revenue by age (black squares with 95 CI, based on local linear regressions with bandwidth 8 months). The red area indicates the mechanical fiscal effect of the reform.

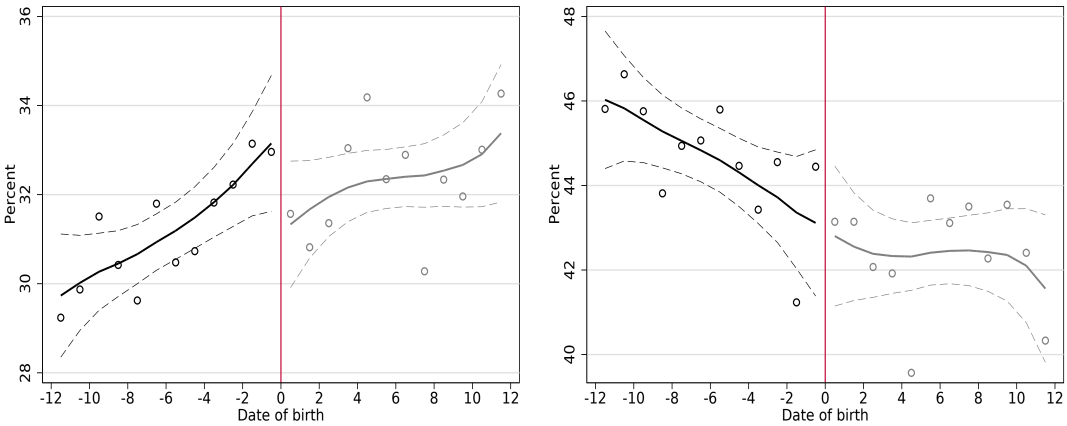
Table 7: Multiplier Benefit Generosity by Discount Rate and Bandwidth

Discount rate:	r=0%	r=2%	r=5%	r=7%
Multiplier Men				
bandwith = 2	2.34	2.62	3.08	3.39
bandwith = 4	1.61	1.75	1.96	2.10
bandwith = 6	1.48	1.59	1.76	1.88
bandwith = 8	1.39	1.48	1.62	1.72
bandwith = 10	1.30	1.37	1.49	1.57
Multiplier Women				
bandwith = 2	1.92	2.15	2.52	2.77
bandwith = 4	1.78	1.98	2.31	2.55
bandwith = 6	1.94	2.19	2.59	2.88
bandwith = 8	1.66	1.84	2.12	2.32
bandwith = 10	1.45	1.56	1.76	1.89

D.3 RDD Figures Women

Figure 40: Retirement and Employment at Age 55: Women

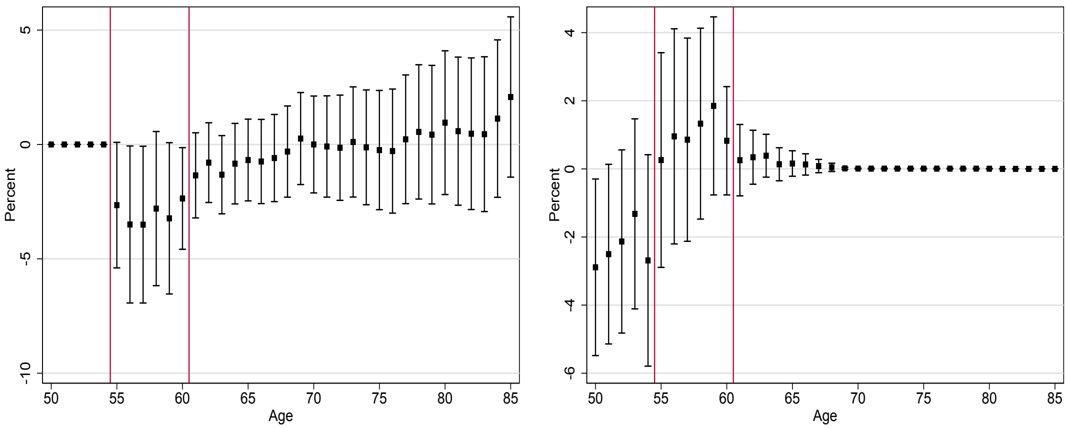
(a) Retirement (OA and DI) (b) Employment



Notes: This figure shows the average retirement and employment rates at age 60. The fitted lines are local linear polynomials with a bandwidth of 8 months.

Figure 41: RD Estimates by Age: Labor Supply Women

(a) Retirement (OA and DI) (b) Employment

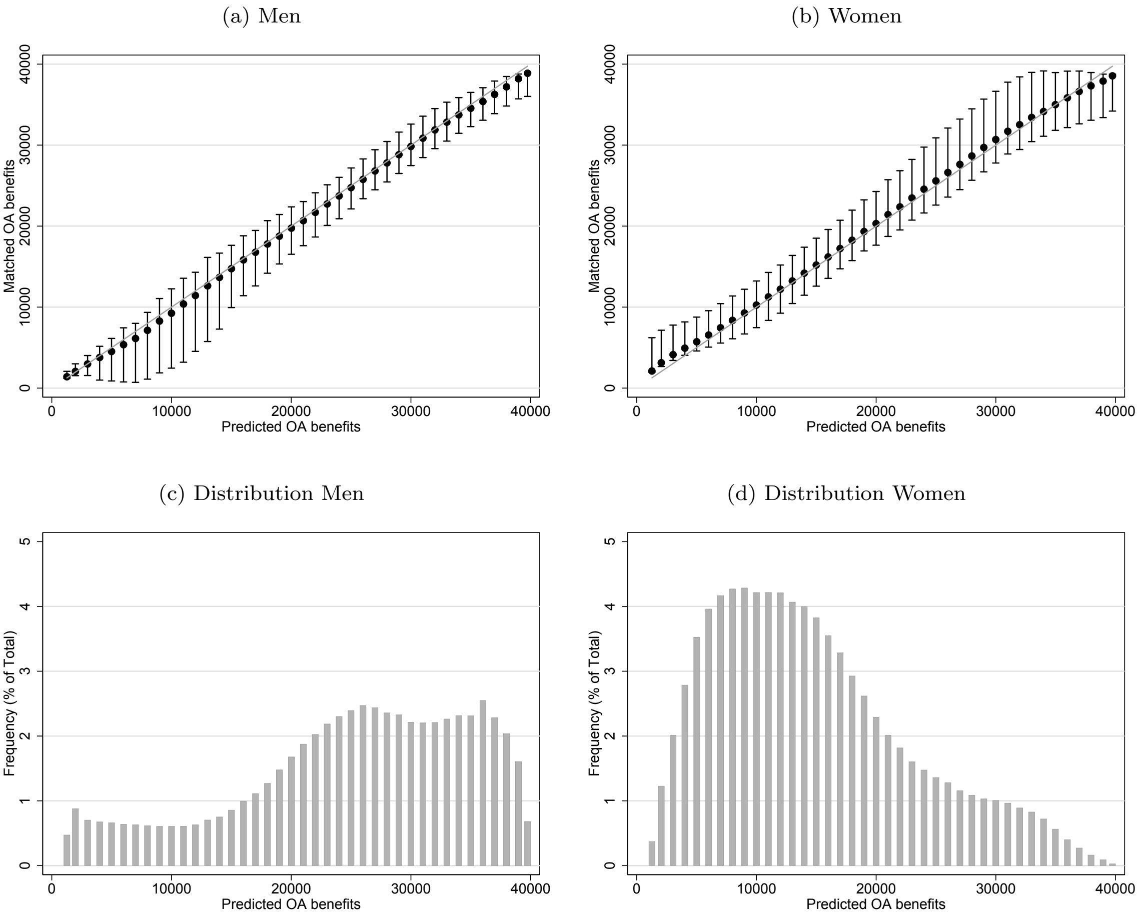


Notes: This figure plots the RD estimates at each age between 50 and 85 based on local linear regressions with a bandwith of 8 months.

E Data Appendix: Pension Benefit Calculation

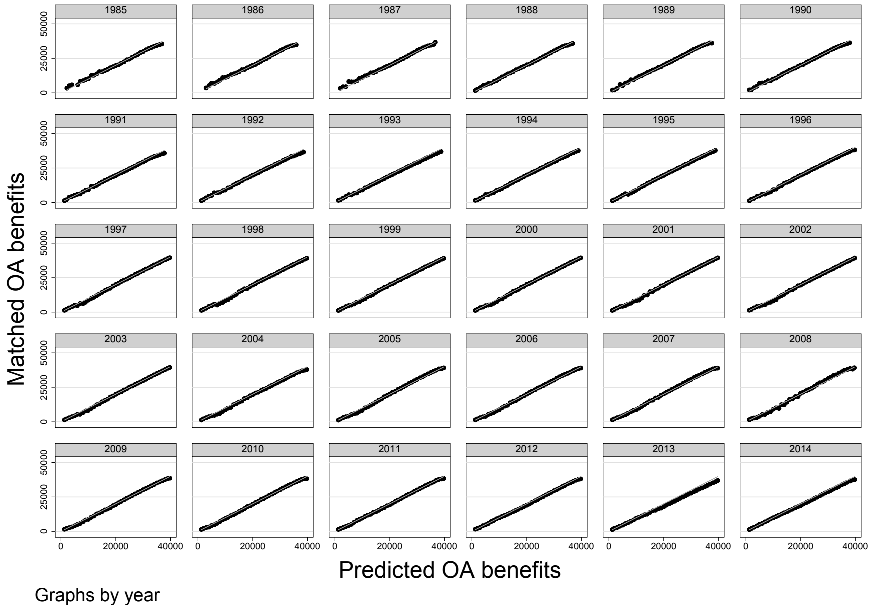
The ASSD provide all information necessary to compute individual old-age and disability pensions. I validate my pension calculations by comparing my predicted old-age pensions with actual old-age pensions for the subsample of retirees who received benefits in 2001 or who retired after 2001, using matched data obtained from the Austrian Social Security Administration. Figure 42 plots mean matched old-age pensions against mean predicted old-age pensions (in 1000-Euro bins), pooling all years from 1985 to 2014. To get a sense of the distribution of predicted benefits I plot the 95th percentile (upper bar) and the 5th percentile (the lower bar), i.e. 90% of all observations lie between the two bars. Figures 43 and 44 plot this for each year separately. Figure 45 shows the same exercise for disability pensions. In all years, actual pensions track predicted pensions very closely.

Figure 42: Imputed and Matched Old-Age Pensions (Years 1985-2014)



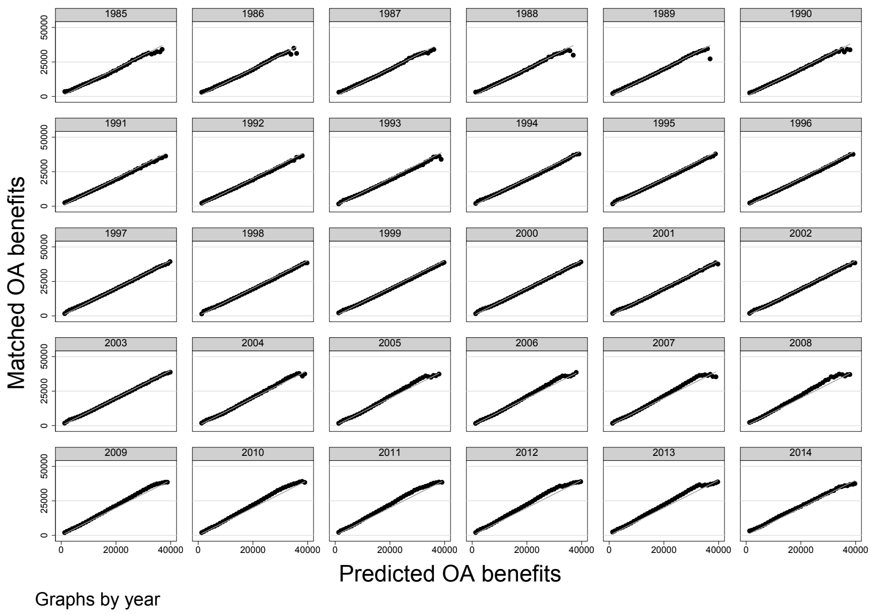
Notes: This figure plots mean matched old-age pensions against mean predicted old-age pensions (in 1000-Euro bins), pooling all years from 1985 to 2014. The bars represent the 95th percentile (upper bar) and the 5th percentile (the lower bar), the dot represents the mean. Panels (c) and (d) show the relative frequency of pension levels for men and women in this subsample.

Figure 43: Imputed and Matched Old-Age Pensions by Year: Men



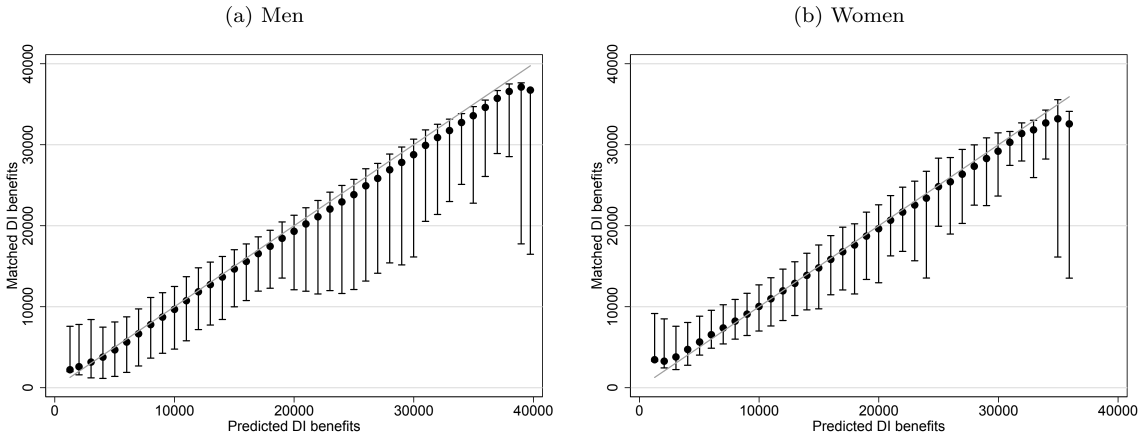
Notes: This figure plots mean matched old-age pensions against mean predicted old-age pensions (in 1000-Euro bins) for each year separately.

Figure 44: Imputed and Matched Old-Age Pensions by Year: Women



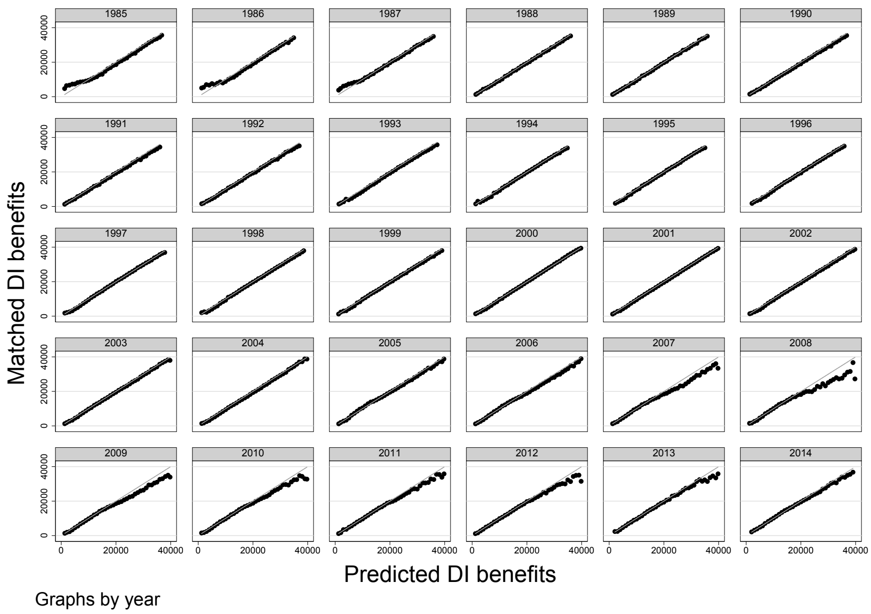
Notes: This figure plots mean matched old-age pensions against mean predicted old-age pensions (in 1000-Euro bins) for each year separately.

Figure 45: Imputed and Matched Disability Pensions (Years 1985-2014)



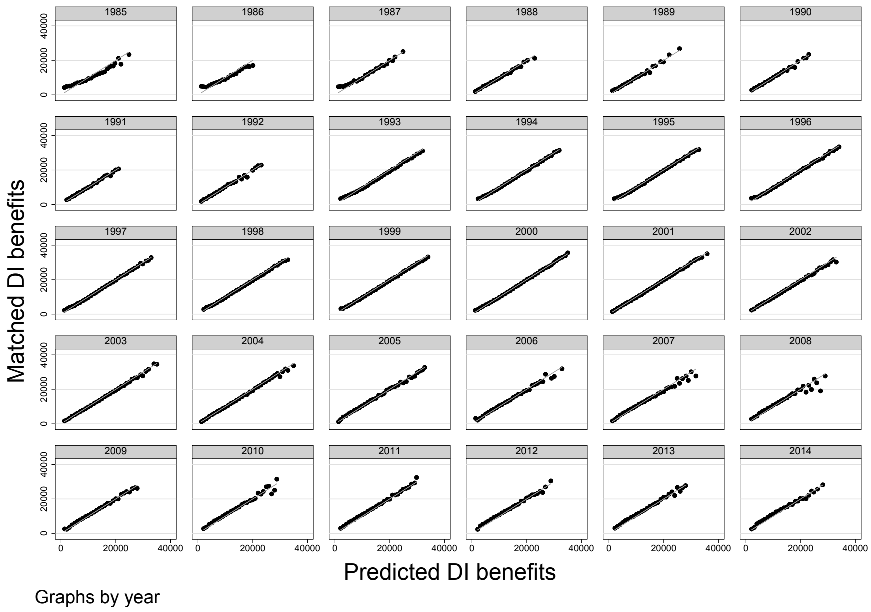
Notes: This figure plots mean matched disability pensions against mean predicted disability pensions (in 1000-Euro bins), pooling all years from 1985 to 2014. The bars represent the 95th percentile (upper bar) and the 5th percentile (the lower bar), the dot represents the mean.

Figure 46: Imputed and Matched Disability Pensions by Year: Men



Notes: This figure plots mean matched disability pensions against mean predicted old-age pensions (in 1000-Euro bins) for each year separately.

Figure 47: Imputed and Matched Disability Pensions by Year: Women



Notes: This figure plots mean matched disability pensions against mean predicted old-age pensions (in 1000-Euro bins) for each year separately.